

**CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL  
PROTEINS**

**Related Applications**

- 5                      This application claims <sup>benefit of</sup> priority to prior filed U.S. Provisional Patent Application Serial No. 60/142764, filed July 8, 1999, and U.S. Provisional Patent Application Serial No. 60/152,318, filed September 3, 1999. The entire contents of both of the aforementioned applications are hereby expressly incorporated herein by reference.

10    **Background of the Invention**

- Certain products and by-products of naturally-occurring metabolic processes in cells have utility in a wide array of industries, including the food, feed, cosmetics, and pharmaceutical industries. These molecules, collectively termed 'fine chemicals', include organic acids, both proteinogenic and non-proteinogenic amino acids, nucleotides and nucleosides, lipids and fatty acids, diols, carbohydrates, aromatic compounds, vitamins and cofactors, and enzymes. Their production is most conveniently performed through the large-scale culture of bacteria developed to produce and secrete large quantities of one or more desired molecules. One particularly useful organism for this purpose is *Corynebacterium glutamicum*, a gram positive, nonpathogenic bacterium. Through strain selection, a number of mutant strains have been developed which produce an array of desirable compounds. However, selection of strains improved for the production of a particular molecule is a time-consuming and difficult process.

25    **Summary of the Invention**

- The invention provides novel bacterial nucleic acid molecules which have a variety of uses. These uses include the identification of microorganisms which can be used to produce fine chemicals, the modulation of fine chemical production in *C. glutamicum* or related bacteria, the typing or identification of *C. glutamicum* or related bacteria, as reference points for mapping the *C. glutamicum* genome, and as markers for transformation. These novel nucleic acid molecules encode proteins, referred to herein as marker and fine chemical production (MCP) proteins.

- C. glutamicum* is a gram positive, aerobic bacterium which is commonly used in industry for the large-scale production of a variety of fine chemicals, and also for the degradation of hydrocarbons (such as in petroleum spills) and for the oxidation of terpenoids. The MCP nucleic acid molecules of the invention, therefore, can be used to

identify microorganisms which can be used to produce fine chemicals, e.g., by

fermentation processes. Modulation of the expression of the MCP nucleic acids of the invention, or modification of the sequence of the MCP nucleic acid molecules of the invention, can be used to modulate the production of one or more fine chemicals from a microorganism (e.g., to improve the yield or production of one or more fine chemicals

5 from a *Corynebacterium* or *Brevibacterium* species).

The MCP nucleic acids of the invention may also be used to identify an organism as being *Corynebacterium glutamicum* or a close relative thereof, or to identify the presence of *C. glutamicum* or a relative thereof in a mixed population of microorganisms. The invention provides the nucleic acid sequences of a number of *C. glutamicum* genes; by probing the extracted genomic DNA of a culture of a unique or mixed population of microorganisms under stringent conditions with a probe spanning a region of a *C. glutamicum* gene which is unique to this organism, one can ascertain whether this organism is present. Although *Corynebacterium glutamicum* itself is nonpathogenic, it is related to species pathogenic in humans, such as *Corynebacterium diphtheriae* (the causative agent of diphtheria); the detection of such organisms is of significant clinical relevance.

The MCP nucleic acid molecules of the invention may also serve as reference points for mapping of the *C. glutamicum* genome, or of genomes of related organisms. Similarly, these molecules, or variants or portions thereof, may serve as markers for genetically engineered *Corynebacterium* or *Brevibacterium* species.

The MCP proteins encoded by the novel nucleic acid molecules of the invention may be involved, for example, in the direct or indirect production of one or more fine chemicals from *C. glutamicum*. The MCP proteins of the invention may also participate in the degradation of hydrocarbons or the oxidation of terpenoids. These proteins may also be utilized for the identification of *Corynebacterium glutamicum* or organisms related to *C. glutamicum*; the presence of an MCP protein specific to *C. glutamicum* and related species in a mixture of proteins may indicate the presence of one of these bacteria in the sample. Further, these MCP proteins may have homologues in plants or animals which are involved in a disease state or condition; these proteins thus may serve as useful pharmaceutical targets for drug screening and the development of therapeutic compounds.

Given the availability of cloning vectors for use in *Corynebacterium glutamicum*, such as those disclosed in Sinskey *et al.*, U.S. Patent No. 4,649,119, and techniques for genetic manipulation of *C. glutamicum* and the related *Brevibacterium* species (e.g., *lactofermentum*) (Yoshihama *et al.*, *J. Bacteriol.* 162: 591-597 (1985); Katsumata *et al.*, *J. Bacteriol.* 159: 306-311 (1984); and Santamaria *et al.*, *J. Gen.*

utilized in the genetic engineering of this organism to modulate the production of one or more fine chemicals. This modulation may be due to a direct effect of manipulation of a gene of the invention, or it may be due to an indirect effect of such manipulation. For example, by modifying the activity of a protein involved in the biosynthesis or

5 degradation of a fine chemical (*i.e.*, through mutagenesis of the corresponding gene), one may directly modulate the ability of the cell to synthesize or to degrade this compound, thereby modulating the yield and/or efficiency of production of the fine chemical. Similarly, by modulating the activity of a protein which regulates a fine chemical metabolic pathway, one may directly influence whether the production of the  
10 desired compound is up- or down-regulated, either of which will modulate the yield or efficiency of production of the fine chemical from the cell.

Indirect modulation of fine chemical production may also result by modifying the activity of a protein of the invention (*i.e.*, by mutagenesis of the corresponding gene) such that the overall ability of the cell to grow and divide or to remain viable and  
15 productive is increased. The production of fine chemicals from *C. glutamicum* is generally accomplished by the large-scale fermentative culture of these microorganisms, conditions which are frequently suboptimal for growth and cell division. By engineering a protein of the invention (*e.g.*, a stress response protein, a cell wall protein, or proteins involved in the metabolism of compounds necessary for cell growth and  
20 division to occur, such as nucleotides and amino acids) such that it is better able to survive, grow, and multiply in such conditions, it may be possible to increase the number and productivity of such engineered *C. glutamicum* cells in large-scale culture, which in turn should result in increased yields and/or efficiency of production of one or more desired fine chemicals. Further, the metabolic pathways of any cell are necessarily  
25 interrelated and coregulated. By altering the activity or regulation of any one metabolic pathway in *C. glutamicum* (*i.e.*, by altering the activity of one of the proteins of the invention which participates in such a pathway), it is possible to concomitantly alter the activity or regulation of other metabolic pathways in this microorganism, which may be directly involved in the synthesis or degradation of a fine chemical.

30 The invention provides novel nucleic acid molecules which encode proteins, referred to herein as MCP proteins, which are capable of, for example, modulating the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of serving as identifying markers for *C. glutamicum* or related organisms. Nucleic acid molecules encoding an MCP protein are referred to herein as  
35 MCP nucleic acid molecules. In a preferred embodiment, the MCP protein is capable of modulating the production or efficiency of production of one or more fine chemicals

organisms. Examples of such proteins include those encoded by the genes set forth in Table 1.

Accordingly, one aspect of the invention pertains to isolated nucleic acid molecules (*e.g.*, cDNAs, DNAs, or RNAs) comprising a nucleotide sequence encoding an MCP protein or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection or amplification of MCP-encoding nucleic acid (*e.g.*, DNA or mRNA). In particularly preferred embodiments, the isolated nucleic acid molecule comprises one of the nucleotide sequences set forth in Appendix A or the coding region or a complement thereof of one of these nucleotide sequences. In other particularly preferred embodiments, the isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes to or is at least about 50%, preferably at least about 60%, more preferably at least about 70%, 80% or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence set forth in Appendix A, or a portion thereof. In other preferred embodiments, the isolated nucleic acid molecule encodes one of the amino acid sequences set forth in Appendix B. The preferred MCP proteins of the present invention also preferably possess at least one of the MCP activities described herein.

In another embodiment, the isolated nucleic acid molecule encodes a protein or portion thereof wherein the protein or portion thereof includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of Appendix B, *e.g.*, sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains an MCP activity. Preferably, the protein or portion thereof encoded by the nucleic acid molecule maintains the ability to modulate the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of serving as an identifying marker for *C. glutamicum* or related organisms. In one embodiment, the protein encoded by the nucleic acid molecule is at least about 50%, preferably at least about 60%, and more preferably at least about 70%, 80%, or 90% and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an amino acid sequence of Appendix B (*e.g.*, an entire amino acid sequence selected from those sequences set forth in Appendix B). In another preferred embodiment, the protein is a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of Appendix B (encoded by an open reading frame shown in Appendix A).

In another preferred embodiment, the isolated nucleic acid molecule is derived from *C. glutamicum* and encodes a protein (*e.g.*, an MCP fusion protein) which includes



the amino acid sequences of Appendix B and is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms, and which  
5 also includes heterologous nucleic acid sequences encoding a heterologous polypeptide or regulatory regions.

In another embodiment, the isolated nucleic acid molecule is at least 15 nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising a nucleotide sequence of Appendix A. Preferably, the isolated  
10 nucleic acid molecule corresponds to a naturally-occurring nucleic acid molecule. More preferably, the isolated nucleic acid encodes a naturally-occurring *C. glutamicum* MCP protein, or a biologically active portion thereof.

Another aspect of the invention pertains to vectors, *e.g.*, recombinant expression vectors, containing the nucleic acid molecules of the invention, and host cells into which  
15 such vectors have been introduced. In one embodiment, such a host cell is used to produce an MCP protein by culturing the host cell in a suitable medium. The MCP protein can then be isolated from the medium or the host cell.

Yet another aspect of the invention pertains to a genetically altered microorganism in which an MCP gene has been introduced or altered. In one  
20 embodiment, the genome of the microorganism has been altered by introduction of a nucleic acid molecule of the invention encoding wild-type or mutated MCP sequence as a transgene. In another embodiment, an endogenous MCP gene within the genome of the microorganism has been altered, *e.g.*, functionally disrupted, by homologous recombination with an altered MCP gene. In another embodiment, an endogenous or  
25 introduced MCP gene in a microorganism has been altered by one or more point mutations, deletions, or inversions, but still encodes a functional MCP protein. In still another embodiment, one or more of the regulatory regions (*e.g.*, a promoter, repressor, or inducer) of an MCP gene in a microorganism has been altered (*e.g.*, by deletion, truncation, inversion, or point mutation) such that the expression of the MCP gene is  
30 modulated. In a preferred embodiment, the microorganism belongs to the genus *Corynebacterium* or *Brevibacterium*, with *Corynebacterium glutamicum* being particularly preferred. In a preferred embodiment, the microorganism is also utilized for the production of a desired compound, such as an amino acid, with lysine being particularly preferred.

35 In another aspect, the invention provides a method of identifying the presence or activity of *Corynebacterium diphtheriae* in a subject. This method includes detection of one or more of the nucleic acid or amino acid sequences of the invention (*e.g.*, the

sequences set forth in Appendix A or Appendix B) in a subject, thereby detecting the presence or activity of *Corynebacterium diphtheriae* in the subject.

Still another aspect of the invention pertains to an isolated MCP protein or a portion, *e.g.*, a biologically active portion, thereof. In a preferred embodiment, the isolated MCP protein or portion thereof is capable of modulating the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of serving as an identifying marker for *C. glutamicum* or related organisms. In another preferred embodiment, the isolated MCP protein or portion thereof is sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains the ability to, for example, modulate the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or to serve as identifying markers for *C. glutamicum* or related organisms.

The invention also provides an isolated preparation of an MCP protein. In preferred embodiments, the MCP protein comprises an amino acid sequence of Appendix B. In another preferred embodiment, the invention pertains to an isolated full length protein which is substantially homologous to an entire amino acid sequence of Appendix B (encoded by an open reading frame set forth in Appendix A). In yet another embodiment, the protein is at least about 50%, preferably at least about 60%, and more preferably at least about 70%, 80%, or 90%, and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an entire amino acid sequence of Appendix B. In other embodiments, the isolated MCP protein comprises an amino acid sequence which is at least about 50% or more homologous to one of the amino acid sequences of Appendix B and is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms.

Alternatively, the isolated MCP protein can comprise an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, or is at least about 50%, preferably at least about 60%, more preferably at least about 70%, 80%, or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous, to a nucleotide sequence of Appendix B. It is also preferred that the preferred forms of MCP proteins also have one or more of the MCP bioactivities described herein.

The MCP polypeptide, or a biologically active portion thereof, can be operatively linked to a non-MCP polypeptide to form a fusion protein. In preferred embodiments, this fusion protein has an activity which differs from that of the MCP protein alone. In other preferred embodiments, this fusion protein is capable of

modulating the yield, production and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of serving as an identifying marker for *C. glutamicum* or related organisms. In particularly preferred embodiments, integration of this fusion protein into a host cell modulates production of a desired compound from the cell.

5 In another aspect, the invention provides methods for screening molecules which modulate the activity of an MCP protein, either by interacting with the protein itself or a substrate or binding partner of the MCP protein, or by modulating the transcription or translation of an MCP nucleic acid molecule of the invention.

Another aspect of the invention pertains to a method for producing a fine  
10 chemical. This method involves the culturing of a cell containing a vector directing the expression of an MCP nucleic acid molecule of the invention, such that a fine chemical is produced. In a preferred embodiment, this method further includes the step of obtaining a cell containing such a vector, in which a cell is transfected with a vector directing the expression of an MCP nucleic acid. In another preferred embodiment, this  
15 method further includes the step of recovering the fine chemical from the culture. In a particularly preferred embodiment, the cell is from the genus *Corynebacterium* or *Brevibacterium*, or is selected from those strains set forth in Table 3.

Another aspect of the invention pertains to methods for modulating production of a molecule from a microorganism. Such methods include contacting the cell with an  
20 agent which modulates MCP protein activity or MCP nucleic acid expression such that a cell associated activity is altered relative to this same activity in the absence of the agent. In a preferred embodiment, the cell is modulated for one or more *C. glutamicum* MCP protein activities, such that the yield, production, and/or efficiency of production of a desired fine chemical by this microorganism is improved. The agent which  
25 modulates MCP protein activity can be an agent which stimulates MCP protein activity or MCP nucleic acid expression. Examples of agents which stimulate MCP protein activity or MCP nucleic acid expression include small molecules, active MCP proteins, and nucleic acids encoding MCP proteins that have been introduced into the cell. Examples of agents which inhibit MCP activity or expression include small molecules  
30 and antisense MCP nucleic acid molecules.

Another aspect of the invention pertains to methods for modulating yields, production, and/or efficiency of production of a desired compound from a cell, involving the introduction of a wild-type or mutant MCP gene into a cell, either maintained on a separate plasmid or integrated into the genome of the host cell. If  
35 integrated into the genome, such integration can be random, or it can take place by homologous recombination such that the native gene is replaced by the introduced copy, causing the production of the desired compound from the cell to be modulated. In a

preferred embodiment, said yields are increased. In another preferred embodiment, said chemical is a fine chemical. In a particularly preferred embodiment, said fine chemical is an amino acid. In especially preferred embodiments, said amino acid is L-lysine.

## 5 **Detailed Description of the Invention**

The present invention provides MCP nucleic acid and protein molecules. These MCP nucleic acid molecules may be utilized in the identification of *Corynebacterium glutamicum* or related organisms, in the mapping of the *C. glutamicum* genome (or a genome of a closely related organism), or in the identification of microorganisms which  
 10 may be used to produce fine chemicals, *e.g.*, by fermentation processes. The proteins encoded by these nucleic acids may be utilized in the direct or indirect modulation of the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, as identifying markers for *C. glutamicum* or related organisms, in the oxidation of terpenoids or the degradation of hydrocarbons, or as targets for the  
 15 development of therapeutic pharmaceutical compounds. Aspects of the invention are further explicated below.

### I. Fine Chemicals

The term 'fine chemical' is art-recognized and includes molecules produced by  
 20 an organism which have applications in various industries, such as, but not limited to, the pharmaceutical, agriculture, and cosmetics industries. Such compounds include organic acids, such as tartaric acid, itaconic acid, and diaminopimelic acid, both proteinogenic and non-proteinogenic amino acids, purine and pyrimidine bases, nucleosides, and nucleotides (as described *e.g.* in Kuninaka, A. (1996) Nucleotides and  
 25 related compounds, p. 561-612, in Biotechnology vol. 6, Rehm *et al.*, eds. VCH: Weinheim, and references contained therein), lipids, both saturated and unsaturated fatty acids (*e.g.*, arachidonic acid), diols (*e.g.*, propane diol, and butane diol), carbohydrates (*e.g.*, hyaluronic acid and trehalose), aromatic compounds (*e.g.*, aromatic amines, vanillin, and indigo), vitamins and cofactors (as described in Ullmann's Encyclopedia of  
 30 Industrial Chemistry, vol. A27, "Vitamins", p. 443-613 (1996) VCH: Weinheim and references therein; and Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research – Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press, (1995)), enzymes,  
 35 polyketides (Cane *et al.* (1998) *Science* 282: 63-68), and all other chemicals described in

0818805086 and references therein. The metabolism and uses of certain of these fine chemicals are further explicated below.

*A. Amino Acid Metabolism and Uses*

5 Amino acids comprise the basic structural units of all proteins, and as such are essential for normal cellular functioning in all organisms. The term "amino acid" is art-recognized. The proteinogenic amino acids, of which there are 20 species, serve as structural units for proteins, in which they are linked by peptide bonds, while the nonproteinogenic amino acids (hundreds of which are known) are not normally found in  
10 proteins (see Ulmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97 VCH: Weinheim (1985)). Amino acids may be in the D- or L- optical configuration, though L-amino acids are generally the only type found in naturally-occurring proteins. Biosynthetic and degradative pathways of each of the 20 proteinogenic amino acids have been well characterized in both prokaryotic and eukaryotic cells (see, for example,  
15 Stryer, L. Biochemistry, 3<sup>rd</sup> edition, pages 578-590 (1988)). The 'essential' amino acids (histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine), so named because they are generally a nutritional requirement due to the complexity of their biosyntheses, are readily converted by simple biosynthetic pathways to the remaining 11 'nonessential' amino acids (alanine, arginine, asparagine, aspartate,  
20 cysteine, glutamate, glutamine, glycine, proline, serine, and tyrosine). Higher animals do retain the ability to synthesize some of these amino acids, but the essential amino acids must be supplied from the diet in order for normal protein synthesis to occur.

Aside from their function in protein biosynthesis, these amino acids are interesting chemicals in their own right, and many have been found to have various  
25 applications in the food, feed, chemical, cosmetics, agriculture, and pharmaceutical industries. Lysine is an important amino acid in the nutrition not only of humans, but also of monogastric animals such as poultry and swine. Glutamate is most commonly used as a flavor additive (mono-sodium glutamate, MSG) and is widely used throughout the food industry, as are aspartate, phenylalanine, glycine, and cysteine. Glycine, L-  
30 methionine and tryptophan are all utilized in the pharmaceutical industry. Glutamine, valine, leucine, isoleucine, histidine, arginine, proline, serine and alanine are of use in both the pharmaceutical and cosmetics industries. Threonine, tryptophan, and D/ L-methionine are common feed additives. (Leuchtenberger, W. (1996) Amino acids - technical production and use, p. 466-502 in Rehm *et al.* (eds.) Biotechnology vol. 6,  
35 chapter 14a, VCH: Weinheim). Additionally, these amino acids have been found to be useful as precursors for the synthesis of synthetic amino acids and proteins, such as N-acetylcysteine, S-carboxymethyl-L-cysteine, (S)-5-hydroxytryptophan, and others

described in Ulmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97, VCH: Weinheim, 1985.

The biosynthesis of these natural amino acids in organisms capable of producing them, such as bacteria, has been well characterized (for review of bacterial amino acid biosynthesis and regulation thereof, see Umbarger, H.E.(1978) *Ann. Rev. Biochem.* 47: 533-606). Glutamate is synthesized by the reductive amination of  $\alpha$ -ketoglutarate, an intermediate in the citric acid cycle. Glutamine, proline, and arginine are each subsequently produced from glutamate. The biosynthesis of serine is a three-step process beginning with 3-phosphoglycerate (an intermediate in glycolysis), and resulting in this amino acid after oxidation, transamination, and hydrolysis steps. Both cysteine and glycine are produced from serine; the former by the condensation of homocysteine with serine, and the latter by the transfer of the side-chain  $\beta$ -carbon atom to tetrahydrofolate, in a reaction catalyzed by serine transhydroxymethylase. Phenylalanine, and tyrosine are synthesized from the glycolytic and pentose phosphate pathway precursors erythrose 4-phosphate and phosphoenolpyruvate in a 9-step biosynthetic pathway that differ only at the final two steps after synthesis of prephenate. Tryptophan is also produced from these two initial molecules, but its synthesis is an 11-step pathway. Tyrosine may also be synthesized from phenylalanine, in a reaction catalyzed by phenylalanine hydroxylase. Alanine, valine, and leucine are all biosynthetic products of pyruvate, the final product of glycolysis. Aspartate is formed from oxaloacetate, an intermediate of the citric acid cycle. Asparagine, methionine, threonine, and lysine are each produced by the conversion of aspartate. Isoleucine is formed from threonine. A complex 9-step pathway results in the production of histidine from 5-phosphoribosyl-1-pyrophosphate, an activated sugar.

Amino acids in excess of the protein synthesis needs of the cell cannot be stored, and are instead degraded to provide intermediates for the major metabolic pathways of the cell (for review see Stryer, L. *Biochemistry* 3<sup>rd</sup> ed. Ch. 21 "Amino Acid Degradation and the Urea Cycle" p. 495-516 (1988)). Although the cell is able to convert unwanted amino acids into useful metabolic intermediates, amino acid production is costly in terms of energy, precursor molecules, and the enzymes necessary to synthesize them. Thus it is not surprising that amino acid biosynthesis is regulated by feedback inhibition, in which the presence of a particular amino acid serves to slow or entirely stop its own production (for overview of feedback mechanisms in amino acid biosynthetic pathways, see Stryer, L. *Biochemistry*, 3<sup>rd</sup> ed. Ch. 24: "Biosynthesis of Amino Acids and Heme" p. 575-600 (1988)). Thus, the output of any particular amino acid is limited by the amount

*B. Vitamin, Cofactor, and Nutraceutical Metabolism and Uses*

Vitamins, cofactors, and nutraceuticals comprise another group of molecules which the higher animals have lost the ability to synthesize and so must ingest, although they are readily synthesized by other organisms such as bacteria. These molecules are

5 either bioactive substances themselves, or are precursors of biologically active substances which may serve as electron carriers or intermediates in a variety of metabolic pathways. Aside from their nutritive value, these compounds also have significant industrial value as coloring agents, antioxidants, and catalysts or other processing aids. (For an overview of the structure, activity, and industrial applications  
10 of these compounds, see, for example, Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613, VCH: Weinheim, 1996.) The term "vitamin" is art-recognized, and includes nutrients which are required by an organism for normal functioning, but which that organism cannot synthesize by itself. The group of vitamins may encompass cofactors and nutraceutical compounds. The language "cofactor"  
15 includes nonproteinaceous compounds required for a normal enzymatic activity to occur. Such compounds may be organic or inorganic; the cofactor molecules of the invention are preferably organic. The term "nutraceutical" includes dietary supplements having health benefits in plants and animals, particularly humans. Examples of such molecules are vitamins, antioxidants, and also certain lipids (*e.g.*, polyunsaturated fatty  
20 acids).

The biosynthesis of these molecules in organisms capable of producing them, such as bacteria, has been largely characterized (Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley  
25 & Sons; Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research – Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press: Champaign, IL X, 374 S).

Thiamin (vitamin B<sub>1</sub>) is produced by the chemical coupling of pyrimidine and  
30 thiazole moieties. Riboflavin (vitamin B<sub>2</sub>) is synthesized from guanosine-5'-triphosphate (GTP) and ribose-5'-phosphate. Riboflavin, in turn, is utilized for the synthesis of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). The family of compounds collectively termed 'vitamin B<sub>6</sub>' (*e.g.*, pyridoxine, pyridoxamine, pyridoxa-5'-phosphate, and the commercially used pyridoxin hydrochloride) are all derivatives of  
35 the common structural unit, 5-hydroxy-6-methylpyridine. Pantothenate (pantothenic acid, (R)-(+)-N-(2,4-dihydroxy-3,3-dimethyl-1-oxobutyl)-β-alanine) can be produced either by chemical synthesis or by fermentation. The final steps in pantothenate

biosynthesis consist of the ATP-driven condensation of  $\beta$ -alanine and pantoic acid. The enzymes responsible for the biosynthesis steps for the conversion to pantoic acid, to  $\beta$ -alanine and for the condensation to panthotenic acid are known. The metabolically active form of pantothenate is Coenzyme A, for which the biosynthesis proceeds in 5

- 5 enzymatic steps. Pantothenate, pyridoxal-5'-phosphate, cysteine and ATP are the precursors of Coenzyme A. These enzymes not only catalyze the formation of panthothante, but also the production of (R)-pantoic acid, (R)-pantolacton, (R)-panthenol (provitamin B<sub>5</sub>), pantetheine (and its derivatives) and coenzyme A.

- 10 Biotin biosynthesis from the precursor molecule pimeloyl-CoA in microorganisms has been studied in detail and several of the genes involved have been identified. Many of the corresponding proteins have been found to also be involved in Fe-cluster synthesis and are members of the nifS class of proteins. Lipoic acid is derived from octanoic acid, and serves as a coenzyme in energy metabolism, where it becomes part of the pyruvate dehydrogenase complex and the  $\alpha$ -ketoglutarate
- 15 dehydrogenase complex. The folates are a group of substances which are all derivatives of folic acid, which is turn is derived from L-glutamic acid, p-amino-benzoic acid and 6-methylpterin. The biosynthesis of folic acid and its derivatives, starting from the metabolism intermediates guanosine-5'-triphosphate (GTP), L-glutamic acid and p-amino-benzoic acid has been studied in detail in certain microorganisms.

- 20 Corrinoids (such as the cobalamines and particularly vitamin B<sub>12</sub>) and porphyrines belong to a group of chemicals characterized by a tetrapyrrole ring system. The biosynthesis of vitamin B<sub>12</sub> is sufficiently complex that it has not yet been completely characterized, but many of the enzymes and substrates involved are now known. Nicotinic acid (nicotinate), and nicotinamide are pyridine derivatives which are
- 25 also termed 'niacin'. Niacin is the precursor of the important coenzymes NAD (nicotinamide adenine dinucleotide) and NADP (nicotinamide adenine dinucleotide phosphate) and their reduced forms.

- The large-scale production of these compounds has largely relied on cell-free chemical syntheses, though some of these chemicals have also been produced by large-
- 30 scale culture of microorganisms, such as riboflavin, Vitamin B<sub>6</sub>, pantothenate, and biotin. Only Vitamin B<sub>12</sub> is produced solely by fermentation, due to the complexity of its synthesis. *In vitro* methodologies require significant inputs of materials and time, often at great cost.

35 *C. Purine, Pyrimidine, Nucleoside and Nucleotide Metabolism and Uses*

Purine and pyrimidine metabolism genes and their corresponding proteins are



“purine” or “pyrimidine” includes the nitrogenous bases which are constituents of nucleic acids, co-enzymes, and nucleotides. The term “nucleotide” includes the basic structural units of nucleic acid molecules, which are comprised of a nitrogenous base, a pentose-sugar (in the case of RNA, the sugar is ribose; in the case of DNA, the sugar is D-deoxyribose), and phosphoric acid. The language “nucleoside” includes molecules which serve as precursors to nucleotides, but which are lacking the phosphoric acid moiety that nucleotides possess. By inhibiting the biosynthesis of these molecules, or their mobilization to form nucleic acid molecules, it is possible to inhibit RNA and DNA synthesis; by inhibiting this activity in a fashion targeted to cancerous cells, the ability of tumor cells to divide and replicate may be inhibited. Additionally, there are nucleotides which may serve as energy stores (e.g., ADP, ATP) or as coenzymes (i.e., FAD and NAD).

Several publications have described the use of these chemicals for these medical indications, by influencing purine and/or pyrimidine metabolism (e.g. Christopherson, R.I. and Lyons, S.D. (1990) “Potent inhibitors of *de novo* pyrimidine and purine biosynthesis as chemotherapeutic agents.” *Med. Res. Reviews* 10: 505-548). Studies of enzymes involved in purine and pyrimidine metabolism have been focused on the development of new drugs which can be used, for example, as immunosuppressants or anti-proliferants (Smith, J.L., (1995) “Enzymes in nucleotide synthesis.” *Curr. Opin. Struct. Biol.* 5: 752-757; (1995) *Biochem Soc. Transact.* 23: 877-902). However, purine and pyrimidine bases, nucleosides and nucleotides have other utilities: as intermediates in the biosynthesis of several fine chemicals (e.g., thiamine, S-adenosyl-methionine, folates, or riboflavin), as energy carriers for the cell (e.g., ATP or GTP), and for chemicals themselves, commonly used as flavor enhancers (e.g., IMP or GMP) or for several medicinal applications (see, for example, Kuninaka, A. (1996) *Nucleotides and Related Compounds in Biotechnology* vol. 6, Rehm *et al.*, eds. VCH: Weinheim, p. 561-612). Also, enzymes involved in purine, pyrimidine, nucleoside, or nucleotide metabolism are increasingly serving as targets against which chemicals for crop protection, including fungicides, herbicides and insecticides, are developed.

The metabolism of these compounds in bacteria has been characterized (for reviews see, for example, Zalkin, H. and Dixon, J.E. (1992) “*de novo* purine nucleotide biosynthesis”, in: *Progress in Nucleic Acid Research and Molecular Biology*, vol. 42, Academic Press:, p. 259-287; and Michal, G. (1999) “Nucleotides and Nucleosides”, Chapter 8 in: *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, Wiley: New York). Purine metabolism has been the subject of intensive research, and is essential to the normal functioning of the cell. Impaired purine metabolism in higher

ribose-5-phosphate, in a series of steps through the intermediate compound inosine-5'-phosphate (IMP), resulting in the production of guanosine-5'-monophosphate (GMP) or adenosine-5'-monophosphate (AMP), from which the triphosphate forms utilized as nucleotides are readily formed. These compounds are also utilized as energy stores, so their degradation provides energy for many different biochemical processes in the cell. Pyrimidine biosynthesis proceeds by the formation of uridine-5'-monophosphate (UMP) from ribose-5-phosphate. UMP, in turn, is converted to cytidine-5'-triphosphate (CTP). The deoxy- forms of all of these nucleotides are produced in a one step reduction reaction from the diphosphate ribose form of the nucleotide to the diphosphate deoxyribose form of the nucleotide. Upon phosphorylation, these molecules are able to participate in DNA synthesis.

#### *D. Trehalose Metabolism and Uses*

Trehalose consists of two glucose molecules, bound in  $\alpha$ ,  $\alpha$ -1,1 linkage. It is commonly used in the food industry as a sweetener, an additive for dried or frozen foods, and in beverages. However, it also has applications in the pharmaceutical, cosmetics and biotechnology industries (see, for example, Nishimoto *et al.*, (1998) U.S. Patent No. 5,759,610; Singer, M.A. and Lindquist, S. (1998) *Trends Biotech.* 16: 460-467; Paiva, C.L.A. and Panek, A.D. (1996) *Biotech. Ann. Rev.* 2: 293-314; and Shiosaka, M. (1997) *J. Japan* 172: 97-102). Trehalose is produced by enzymes from many microorganisms and is naturally released into the surrounding medium, from which it can be collected using methods known in the art.

## II. Elements and Methods of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as MCP nucleic acid molecules. These MCP nucleic acid molecules are useful not only for the identification of *C. glutamicum* or related bacterial species, but also as markers for the mapping of the *C. glutamicum* genome and in the identification of bacteria useful for the production of fine chemicals by, *e.g.*, fermentative processes. The present invention is also based, at least in part, on the MCP protein molecules encoded by these MCP nucleic acid molecules. These MCP proteins are capable of modulating the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, of serving as identifying markers for *C. glutamicum* or related organisms, of degrading hydrocarbons, and of serving as targets for the development of therapeutic pharmaceutical compounds. In one embodiment, the MCP molecules of the invention directly or indirectly participate in one or more fine chemical metabolic pathways in *C. glutamicum*. In a preferred embodiment, the activity

of the MCP molecules of the invention to indirectly or directly participate in such metabolic pathways has an impact on the production of a desired fine chemical by this microorganism. In a particularly preferred embodiment, the MCP molecules of the invention are modulated in activity, such that the *C. glutamicum* metabolic pathways in

5 which the MCP proteins of the invention participate are modulated in efficiency or output, which either directly or indirectly modulates the production or efficiency of production of a desired fine chemical by *C. glutamicum*.

The language, "MCP protein" or "MCP polypeptide" includes proteins which are able to modulate the yield, production, and/or efficiency of production of one or more  
10 fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target protein for drug screening or design, or to serve as identifying markers for *C. glutamicum* or related organisms. Examples of MCP proteins include those encoded by the MCP genes set forth in Table 1 and Appendix A. The terms "MCP gene" or "MCP nucleic acid sequence" include nucleic acid sequences encoding an  
15 MCP protein, which consist of a coding region and also corresponding untranslated 5' and 3' sequence regions. Examples of MCP genes include those set forth in Table 1. The terms "production" or "productivity" are art-recognized and include the concentration of the fermentation product (for example, the desired fine chemical) formed within a given time and a given fermentation volume (e.g., kg product per hour per liter). The term  
20 "efficiency of production" includes the time required for a particular level of production to be achieved (for example, how long it takes for the cell to attain a particular rate of output of a fine chemical). The term "yield" or "product/carbon yield" is art-recognized and includes the efficiency of the conversion of the carbon source into the product (i.e., fine chemical). This is generally written as, for example, kg product per kg carbon  
25 source. By increasing the yield or production of the compound, the quantity of recovered molecules, or of useful recovered molecules of that compound in a given amount of culture over a given amount of time is increased. The terms "biosynthesis" or a "biosynthetic pathway" are art-recognized and include the synthesis of a compound, preferably an organic compound, by a cell from intermediate compounds in what may  
30 be a multistep and highly regulated process. The terms "degradation" or a "degradation pathway" are art-recognized and include the breakdown of a compound, preferably an organic compound, by a cell to degradation products (generally speaking, smaller or less complex molecules) in what may be a multistep and highly regulated process. The language "metabolism" is art-recognized and includes the totality of the biochemical  
35 reactions that take place in an organism. The metabolism of a particular compound, then, (e.g., the metabolism of an amino acid such as glycine) comprises the overall

biosynthetic, modification, and degradation pathways in the cell related to this compound.

In another embodiment, the MCP molecules of the invention are capable of modulating the production of a desired molecule, such as a fine chemical, in a

5 microorganism such as *C. glutamicum*, either directly or indirectly. Using recombinant genetic techniques, one or more of the MCP proteins of the invention may be manipulated such that its function is modulated. Such modulation of function may result in the modulation of the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*.

10 For example, by modifying the activity of a protein involved in the biosynthesis or degradation of a fine chemical (*i.e.*, through mutagenesis of the corresponding gene), one may directly modulate the ability of the cell to synthesize or to degrade this compound, thereby modulating the yield and/or efficiency of production of the fine chemical. Similarly, by modulating the activity of a protein which regulates a fine  
15 chemical metabolic pathway, one may directly influence whether the production of the desired compound is up- or down-regulated, either of which will modulate the yield or efficiency of production of the fine chemical from the cell.

Indirect modulation of fine chemical production may also result by modifying the activity of a protein of the invention (*i.e.*, by mutagenesis of the corresponding gene)  
20 such that the overall ability of the cell to grow and divide or to remain viable and productive is increased. The production of fine chemicals from *C. glutamicum* is generally accomplished by the large-scale fermentative culture of these microorganisms, conditions which are frequently suboptimal for growth and cell division. By engineering a protein of the invention (*e.g.*, a stress response protein, a cell wall protein,  
25 or proteins involved in the metabolism of compounds necessary for cell growth and division to occur, such as nucleotides and amino acids) such that it is better able to survive, grow, and multiply in such conditions, it may be possible to increase the number and productivity of such engineered *C. glutamicum* cells in large-scale culture, which in turn should result in increased yields and/or efficiency of production of one or  
30 more desired fine chemicals. Further, the metabolic pathways of any cell are necessarily interrelated and coregulated. By altering the activity or regulation of any one metabolic pathway in *C. glutamicum* (*i.e.*, by altering the activity of one of the proteins of the invention which participates in such a pathway), it is possible to concomitantly alter the activity or regulation of other metabolic pathways in this microorganism, which may be  
35 directly involved in the synthesis or degradation of a fine chemical.

Culture Collection, given designation ATCC 13032. The nucleotide sequences of the isolated *C. glutamicum* MCP nucleic acid molecules and the predicted amino acid sequences of the *C. glutamicum* MCP proteins are shown in Appendices A and B, respectively. Computational analyses were performed which classified and/or identified

- 5 many of these nucleotide sequences as sequences having homology to *E. coli* or *Bacillus subtilis* genes.

- The present invention also pertains to proteins which have an amino acid sequence which is substantially homologous to an amino acid sequence of Appendix B. As used herein, a protein which has an amino acid sequence which is substantially
- 10 homologous to a selected amino acid sequence is least about 50% homologous to the selected amino acid sequence, *e.g.*, the entire selected amino acid sequence. A protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence can also be least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, or 90-95%, and most preferably at
- 15 least about 96%, 97%, 98%, 99% or more homologous to the selected amino acid sequence.

- The MCP protein or a biologically active portion or fragment thereof of the invention is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize
- 20 terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms.

Various aspects of the invention are described in further detail in the following subsections:

25 *A. Isolated Nucleic Acid Molecules*

- One aspect of the invention pertains to isolated nucleic acid molecules that encode MCP polypeptides or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes or primers for the identification or amplification of MCP-encoding nucleic acid (*e.g.*, MCP DNA). These nucleic acid
- 30 molecules may be used to identify *C. glutamicum* or related organisms, to map the genome of *C. glutamicum* or closely related bacteria, or to identify microorganisms useful for the production of fine chemicals, *e.g.*, by fermentative processes. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or
- 35 RNA generated using nucleotide analogs. This term also encompasses untranslated sequence located at both the 3' and 5' ends of the coding region of the gene: at least about 100 nucleotides of sequence upstream from the 5' end of the coding region and at

least about 20 nucleotides of sequence downstream from the 3' end of the coding region of the gene. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of

- 5 the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated MCP nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide
- 10 sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g., a *C. glutamicum* cell). Moreover, an "isolated" nucleic acid molecule, such as a DNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized.
- 15 A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having a nucleotide sequence of Appendix A, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, a *C. glutamicum* MCP DNA can be isolated from a *C. glutamicum* library using all or portion of one of the sequences of Appendix A as a hybridization probe and
- 20 standard hybridization techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Moreover, a nucleic acid molecule encompassing all or a portion of one of the sequences of Appendix A can be isolated by the polymerase chain reaction using oligonucleotide
- 25 primers designed based upon this sequence (*e.g.*, a nucleic acid molecule encompassing all or a portion of one of the sequences of Appendix A can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this same sequence of Appendix A). For example, mRNA can be isolated from normal endothelial cells (*e.g.*, by the guanidinium-thiocyanate extraction procedure of Chirgwin *et al.* (1979)
- 30 *Biochemistry* 18: 5294-5299) and DNA can be prepared using reverse transcriptase (*e.g.*, Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from Seikagaku America, Inc., St. Petersburg, FL) and random polynucleotide primers or oligonucleotide primers based upon one of the nucleotide sequences shown in Appendix A. Synthetic oligonucleotide primers for
- 35 polymerase chain reaction amplification can be designed based upon one of the nucleotide sequences shown in Appendix A. A nucleic acid of the invention can be amplified using cDNA or, alternatively, genomic DNA, as a template and appropriate

oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to an MCP nucleotide sequence can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises one of the nucleotide sequences shown in Appendix A. The sequences of Appendix A correspond to the *Corynebacterium glutamicum* MCP DNAs of the invention. This cDNA comprises sequences encoding MCP proteins (i.e., the "coding region", indicated in each sequence in Appendix A), as well as 5' untranslated sequences and 3' untranslated sequences, also indicated in Appendix A. Alternatively, the nucleic acid molecule can comprise only the coding region of any of the sequences in Appendix A.

For the purposes of this application, it will be understood that each of the sequences set forth in Appendix A has an identifying RXA or RXN number having the designation "RXA" or "RXN" followed by 5 digits (i.e., RXA00003 or RXN00022). Each of these sequences comprises up to three parts: a 5' upstream region, a coding region, and a downstream region. Each of these three regions is identified by the same RXA or RXN designation to eliminate confusion. The recitation "one of the sequences in Appendix A", then, refers to any of the sequences in Appendix A, which may be distinguished by their differing RXA or RXN designations. The coding region of each of these sequences is translated into a corresponding amino acid sequence, which is set forth in Appendix B. The sequences of Appendix B are identified by the same RXA or RXN designations as Appendix A, such that they can be readily correlated. For example, the amino acid sequence in Appendix B designated RXA00003 is a translation of the coding region of the nucleotide sequence of nucleic acid molecule RXA00003 in Appendix A, and the amino acid sequence in Appendix B designated RXN00022 is a translation of the coding region of the nucleotide sequence of nucleic acid molecule RXN00022 in Appendix A. Each of the RXA and RXN nucleotide and amino acid sequences of the invention has also been assigned a SEQ ID NO, as indicated in Table 1.

Several of the genes of the invention are "F-designated genes". An F-designated gene includes those genes set forth in Table 1 which have an 'F' in front of the RXA designation. For example, SEQ ID NO:3, designated, as indicated on Table 1, as "F RXA01638", is an F-designated gene, as are SEQ ID NOs: 5, 9, and 11 (designated on Table 1 as "F RXA01639", "F RXA01590", and "F RXA01542", respectively).

In one embodiment, the nucleic acid molecules of the present invention are not

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of one of the nucleotide sequences shown in Appendix A, or a portion thereof. A nucleic acid molecule which is complementary to one of the nucleotide sequences shown in

- 5 Appendix A is one which is sufficiently complementary to one of the nucleotide sequences shown in Appendix A such that it can hybridize to one of the nucleotide sequences shown in Appendix A, thereby forming a stable duplex.

- In still another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which is at least about 50%, 51%, 52%, 53%,  
10 54%, 55%, 56%, 57%, 58%, 59%, or 60%, preferably at least about 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence shown  
15 in Appendix A, or a portion thereof. Ranges and identity values intermediate to the above-recited ranges, (e.g., 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended to be included. In an additional preferred embodiment, an isolated nucleic acid  
20 molecule of the invention comprises a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to one of the nucleotide sequences shown in Appendix A, or a portion thereof.

- Moreover, the nucleic acid molecule of the invention can comprise only a portion of the coding region of one of the sequences in Appendix A, for example a  
25 fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of an MCP protein. The nucleotide sequences determined from the cloning of the MCP genes from *C. glutamicum* allows for the generation of probes and primers designed for use in identifying and/or cloning MCP homologues in other cell types and organisms, as well as MCP homologues from other *Corynebacteria* or related  
30 species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 40, 50 or 75 consecutive nucleotides of a sense strand of one of the sequences set forth in Appendix A, an anti-sense sequence of one of the sequences set forth in  
35 Appendix A, or naturally occurring mutants thereof. Primers based on a nucleotide sequence of Appendix A can be used in PCR reactions to clone MCP homologues. Probes based on the MCP nucleotide sequences can be used to detect transcripts or



genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.* the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells

- 5 which misexpress an MCP protein, such as by measuring a level of an MCP-encoding nucleic acid in a sample of cells, *e.g.*, detecting MCP mRNA levels or determining whether a genomic MCP gene has been mutated or deleted.

- In one embodiment, the nucleic acid molecule of the invention encodes a protein or portion thereof which includes an amino acid sequence which is sufficiently
- 10 homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains the ability to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. As used herein, the
- 15 language "sufficiently homologous" refers to proteins or portions thereof which have amino acid sequences which include a minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain as an amino acid residue in one of the sequences of Appendix B) amino acid residues to an amino acid sequence of Appendix B such that the protein or portion thereof is able to modulate the yield,
- 20 production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. Examples of such activities are also described herein. Thus, "the function of an MCP protein" contributes to the overall regulation of one or more fine chemical
- 25 metabolic pathways, or to the degradation of a hydrocarbon, or to the oxidation of a terpenoid.

- In another embodiment, the protein is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino
- 30 acid sequence of Appendix B.

- Portions of proteins encoded by the MCP nucleic acid molecules of the invention are preferably biologically active portions of one of the MCP proteins. As used herein, the term "biologically active portion of an MCP protein" is intended to include a portion, *e.g.*, a domain/motif, of an MCP protein that modulates the yield, production, and/or
- 35 efficiency of production of one or more fine chemicals from *C. glutamicum*, that degrades hydrocarbons, that oxidizes terpenoids, that may serve as a target for drug

organisms. To determine whether an MCP protein or a biologically active portion thereof can modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, can degrade hydrocarbons, or can oxidize terpenoids, an assay of activity may be performed. Such assay methods are well known

5 to those of ordinary skill in the art, as detailed in Example 8 of the Exemplification.

Additional nucleic acid fragments encoding biologically active portions of an MCP protein can be prepared by isolating a portion of one of the sequences in Appendix B, expressing the encoded portion of the MCP protein or peptide (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the MCP protein  
10 or peptide.

The invention further encompasses nucleic acid molecules that differ from one of the nucleotide sequences shown in Appendix A (and portions thereof) due to degeneracy of the genetic code and thus encode the same MCP protein as that encoded by the nucleotide sequences shown in Appendix A. In another embodiment, an isolated nucleic  
15 acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in Appendix B. In a still further embodiment, the nucleic acid molecule of the invention encodes a full length *C. glutamicum* protein which is substantially homologous to an amino acid sequence of Appendix B (encoded by an open reading frame shown in Appendix A).

20 It will be understood by one of ordinary skill in the art that in one embodiment the sequences of the invention are not meant to include the sequences of the prior art, such as those Genbank sequences set forth in Tables 2 or 4 which were available prior to the present invention. In one embodiment, the invention includes nucleotide and amino acid sequences having a percent identity to a nucleotide or amino acid sequence of the  
25 invention which is greater than that of a sequence of the prior art (e.g., a Genbank sequence (or the protein encoded by such a sequence) set forth in Tables 2 or 4). For example, the invention includes a nucleotide sequence which is greater than and/or at least 39% identical to the nucleotide sequence designated RXA00008 (SEQ ID NO:1549), a nucleotide sequence which is greater than and/or at least 42% identical to  
30 the nucleotide sequence designated RXA00059 (SEQ ID NO:1571), and a nucleotide sequence which is greater than and/or at least 39% identical to the nucleotide sequence designated RXA00096 (SEQ ID NO:93). One of ordinary skill in the art would be able to calculate the lower threshold of percent identity for any given sequence of the invention by examining the GAP-calculated percent identity scores set forth in Table 4  
35 for each of the three top hits for the given sequence, and by subtracting the highest GAP-calculated percent identity from 100 percent. One of ordinary skill in the art will

greater than the lower threshold so calculated (e.g., at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, or 60%, preferably at least about 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more identical) are also encompassed by the invention.

In addition to the *C. glutamicum* MCP nucleotide sequences shown in Appendix A, it will be appreciated by those of ordinary skill in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of MCP proteins may exist within a population (e.g., the *C. glutamicum* population). Such genetic polymorphism in the MCP gene may exist among individuals within a population due to natural variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding an MCP protein, preferably a *C. glutamicum* MCP protein. Such natural variations can typically result in 1-5% variance in the nucleotide sequence of the MCP gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in MCP that are the result of natural variation and that do not alter the functional activity of MCP proteins are intended to be within the scope of the invention.

Nucleic acid molecules corresponding to natural variants and non-*C. glutamicum* homologues of the *C. glutamicum* MCP DNA of the invention can be isolated based on their homology to the *C. glutamicum* MCP nucleic acid disclosed herein using the *C. glutamicum* DNA, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising a nucleotide sequence of Appendix A. In other embodiments, the nucleic acid is at least 30, 50, 100, 250 or more nucleotides in length. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 65%, more preferably at least about 70%, and even more preferably at least about 75% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those of ordinary skill in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC)

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to a sequence of Appendix A corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature

5 (e.g., encodes a natural protein). In one embodiment, the nucleic acid encodes a natural *C. glutamicum* MCP protein.

In addition to naturally-occurring variants of the MCP sequence that may exist in the population, one of ordinary skill in the art will further appreciate that changes can be introduced by mutation into a nucleotide sequence of Appendix A, thereby leading to  
10 changes in the amino acid sequence of the encoded MCP protein, without altering the functional ability of the MCP protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in a sequence of Appendix A. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of one of the MCP proteins (Appendix B) without  
15 altering the activity of said MCP protein, whereas an "essential" amino acid residue is required for MCP protein activity. Other amino acid residues, however, (e.g., those that are not conserved or only semi-conserved in the domain having MCP activity) may not be essential for activity and thus are likely to be amenable to alteration without altering MCP activity.

20 Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding MCP proteins that contain changes in amino acid residues that are not essential for MCP activity. Such MCP proteins differ in amino acid sequence from a sequence contained in Appendix B yet retain at least one of the MCP activities described herein. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence  
25 encoding a protein, wherein the protein comprises an amino acid sequence at least about 50% homologous to an amino acid sequence of Appendix B and is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related  
30 organisms. Preferably, the protein encoded by the nucleic acid molecule is at least about 50-60% homologous to one of the sequences in Appendix B, more preferably at least about 60-70% homologous to one of the sequences in Appendix B, even more preferably at least about 70-80%, 80-90%, 90-95% homologous to one of the sequences in Appendix B, and most preferably at least about 96%, 97%, 98%, or 99% homologous  
35 to one of the sequences in Appendix B.

The homology between two amino acid sequences (e.g., one of

sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in one sequence (*e.g.*, one of the sequences of Appendix B) is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence (*e.g.*, a mutant form of the sequence selected from Appendix B), then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity"). The percent homology between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % homology = # of identical positions/total # of positions x 100).

An isolated nucleic acid molecule encoding an MCP protein homologous to a protein sequence of Appendix B can be created by introducing one or more nucleotide substitutions, additions or deletions into a nucleotide sequence of Appendix A such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into one of the sequences of Appendix A by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in an MCP protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an MCP coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for an MCP activity described herein to identify mutants that retain MCP activity. Following mutagenesis of one of the sequences of Appendix A, the encoded protein can be expressed recombinantly and the activity of the

In addition to the nucleic acid molecules encoding MCP proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*,

5 complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire MCP coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the  
10 coding strand of a nucleotide sequence encoding an MCP protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the entire coding region of SEQ ID NO. 1 (RXN01638) comprises nucleotides 1 to 900). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a  
15 nucleotide sequence encoding MCP. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding MCP disclosed herein (*e.g.*, the sequences set forth in Appendix A), antisense nucleic acids of the invention can be  
20 designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of MCP mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of MCP mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of MCP mRNA. An  
25 antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed by chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified  
30 nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense  
35 nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-

methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a cell or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an MCP protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. The antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecule to a peptide or an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong prokaryotic, viral, or eukaryotic promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an  $\alpha$ -anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-O-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they

(described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave MCP mRNA transcripts to thereby inhibit translation of MCP mRNA. A ribozyme having specificity for an MCP-encoding nucleic acid can be designed based upon the nucleotide sequence of an MCP-DNA disclosed herein (*i.e.*,

- 5 SEQ ID NO. 1 (RXN01368) in Appendix A). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an MCP-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071 and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, MCP mRNA can be used to select a catalytic RNA  
10 having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

- Alternatively, MCP gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of an MCP nucleotide sequence (*e.g.*, an MCP promoter and/or enhancers) to form triple helical structures that prevent  
15 transcription of an MCP gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

#### *B. Recombinant Expression Vectors and Host Cells*

- 20 Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an MCP protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA  
25 segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host  
30 cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can  
35 be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors,



such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, repressor binding sites, activator binding sites, enhancer regions and other expression control elements (*e.g.*, terminators, other elements of mRNA secondary structure, or polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells. Preferred regulatory sequences are, for example, promoters such as *cos*-, *tac*-, *trp*-, *tet*-, *trp-tet*-, *lpp*-, *lac*-, *lpp-lac*-, *lacI<sup>q</sup>*-, *T7*-, *T5*-, *T3*-, *gal*-, *trc*-, *ara*-, *SP6*-, *arny*, *SPO2*,  $\lambda$ -*P<sub>R</sub>*- or  $\lambda$  *P<sub>L</sub>*, which are used preferably in bacteria. Additional regulatory sequences are, for example, promoters from yeasts and fungi, such as *ADC1*, *MF $\alpha$* , *AC*, *P-60*, *CYC1*, *GAPDH*, *TEF*, *rp28*, *ADH*, promoters from plants such as *CaMV/35S*, *SSU*, *OCS*, *lib4*, *usp*, *STLS1*, *B33*, *nos* or *ubiquitin*- or *phaseolin*- promoters. It is also possible to use artificial promoters. It will be appreciated by those of ordinary skill in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, MCP proteins, mutant forms of MCP proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of MCP proteins in prokaryotic or eukaryotic cells. For example, MCP genes can be expressed in bacterial cells such as *C. glutamicum*, insect cells (using baculovirus expression vectors), yeast and other fungal cells (see Romanos, M.A. *et al.* (1992) "Foreign gene expression in yeast: a review", *Yeast* 8: 423-488; van den Hondel, C.A.M.J. *et al.* (1991) "Heterologous gene expression in filamentous fungi" in: More

Gene Manipulations in Fungi, J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic Press: San Diego; and van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, Peberdy, J.F. *et al.*, eds., p. 1-28, Cambridge University Press: Cambridge),  
 5 algae and multicellular plant cells (see Schmidt, R. and Willmitzer, L. (1988) High efficiency *Agrobacterium tumefaciens* -mediated transformation of *Arabidopsis thaliana* leaf and cotyledon explants" *Plant Cell Rep.*: 583-586), or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the  
 10 recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein  
 15 encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion  
 20 expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly,  
 25 MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. In one embodiment, the coding sequence of the MCP protein is cloned into a pGEX expression vector to create a vector encoding a fusion protein comprising, from the N-terminus to the C-terminus, GST-thrombin cleavage site-X protein. The fusion  
 30 protein can be purified by affinity chromatography using glutathione-agarose resin. Recombinant MCP protein unfused to GST can be recovered by cleavage of the fusion protein with thrombin.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315), pLG338, pACYC184, pBR322, pUC18,  
 35 pUC19, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III113-B1,  $\lambda$ gt11, pBdCl, and pET 11d (Studier *et al.*, *Gene Expression Technology:*

Pouwels *et al.*, eds. (1985) Cloning Vectors. Elsevier: New York ISBN 0 444 904018). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by

5 a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident  $\lambda$  prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter. For transformation of other varieties of bacteria, appropriate vectors may be selected. For example, the plasmids pIJ101, pIJ364, pIJ702 and pIJ361 are known to be useful in transforming

10 *Streptomyces*, while plasmids pUB110, pC194, or pBD214 are suited for transformation of *Bacillus* species. Several plasmids of use in the transfer of genetic information into *Corynebacterium* include pHM1519, pBL1, pSA77, or pAJ667 (Pouwels *et al.*, eds. (1985) Cloning Vectors. Elsevier: New York ISBN 0 444 904018).

One strategy to maximize recombinant protein expression is to express the

15 protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those

20 preferentially utilized in the bacterium chosen for expression, such as *C. glutamicum* (Wada *et al.* (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the MCP protein expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1

25 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), 2  $\mu$ , pAG-1, Yep6, Yep13, pEMBLYe23, pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Vectors and methods for the construction of vectors appropriate for use in other fungi, such as the filamentous fungi, include those detailed in: van den Hondel, C.A.M.J.J. & Punt, P.J.

30 (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, J.F. Peberdy, *et al.*, eds., p. 1-28, Cambridge University Press: Cambridge, and Pouwels *et al.*, eds. (1985) Cloning Vectors. Elsevier: New York (ISBN 0 444 904018).

Alternatively, the MCP proteins of the invention can be expressed in insect cells

35 using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.* Sf9 cells) include the pAc series (Smith *et al.*

(1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In another embodiment, the MCP proteins of the invention may be expressed in unicellular plant cells (such as algae) or in plant cells from higher plants (e.g., the  
5 spermatophytes, such as crop plants). Examples of plant expression vectors include those detailed in: Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20: 1195-1197; and Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation", *Nucl. Acid. Res.* 12: 8711-8721, and include pLGV23, pGHIac+,  
10 pBIN19, pAK2004, and pDH51 (Pouwels *et al.*, eds. (1985) *Cloning Vectors*. Elsevier: New York ISBN 0 444 904018).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC  
15 (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F.,  
20 and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type  
25 (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and  
30 Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and  
35 European Application Publication No. 264,166). Developmentally-regulated promoters

*Science* 249:374-379) and the  $\alpha$ -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to MCP mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.* (1986) "Antisense RNA as a molecular tool for genetic analysis", *Reviews - Trends in Genetics*, Vol. 1(1).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, an MCP protein can be expressed in bacterial cells such as *C. glutamicum*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those of ordinary skill in the art. Microorganisms related to *Corynebacterium glutamicum* which may be conveniently used as host cells for the nucleic acid and protein molecules of the invention are set forth in Table 3.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation", "transfection", "conjugation" and "transduction" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, linear DNA or RNA (*e.g.*, a linearized vector or a gene construct alone without a vector) or

transposon or other DNA) into a host cell, including using natural competence, chemical mediated transfer, calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular*

- 5 *Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these

10 integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding an MCP protein or can be

15 introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by, for example, drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

To create a homologous recombinant microorganism, a vector is prepared which contains at least a portion of an MCP gene into which a deletion, addition or substitution

20 has been introduced to thereby alter, *e.g.*, functionally disrupt, the MCP gene. Preferably, this MCP gene is a *Corynebacterium glutamicum* MCP gene, but it can be a homologue from a related bacterium or even from a mammalian, yeast, or insect source. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous MCP gene is functionally disrupted (*i.e.*, no longer

25 encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous MCP gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous MCP protein). In the homologous recombination vector, the altered portion

30 of the MCP gene is flanked at its 5' and 3' ends by additional nucleic acid of the MCP gene to allow for homologous recombination to occur between the exogenous MCP gene carried by the vector and an endogenous MCP gene in a microorganism. The additional flanking MCP nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, less than one kilobase of flanking

35 DNA (both at the 5' and 3' ends) is included in the vector (see *e.g.*, Thomas, K.R., and Capecchi, M.R. (1987) *Cell* 51: 503 for a description of homologous recombination vectors). The vector is introduced into a microorganism (*e.g.*, by electroporation) and

cells in which the introduced MCP gene has homologously recombined with the endogenous MCP gene are selected, using art-known techniques.

In another embodiment, recombinant microorganisms can be produced which contain selected systems which allow for regulated expression of the introduced gene.

- 5 For example, inclusion of an MCP gene on a vector placing it under control of the lac operon permits expression of the MCP gene in the presence of IPTG. Such regulatory systems are well known in the art.

- 10 In another embodiment, an endogenous MCP gene in a host cell is disrupted (*e.g.*, by homologous recombination or other genetic means known in the art) such that expression of its protein product does not occur. In another embodiment, an endogenous or introduced MCP gene in a host cell has been altered by one or more point mutations, deletions, or inversions, but still encodes a functional MCP protein. In still another embodiment, one or more of the regulatory regions (*e.g.*, a promoter, repressor, or inducer) of an MCP gene in a microorganism has been altered (*e.g.*, by deletion,
- 15 truncation, inversion, or point mutation) such that the expression of the MCP gene is modulated. One of ordinary skill in the art will appreciate that host cells containing more than one of the described MCP gene and protein modifications may be readily produced using the methods of the invention, and are meant to be included in the present invention.

- 20 A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) an MCP protein. Accordingly, the invention further provides methods for producing MCP proteins using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding an MCP protein has
- 25 been introduced, or into which genome has been introduced a gene encoding a wild-type or altered MCP protein) in a suitable medium until MCP protein is produced. In another embodiment, the method further comprises isolating MCP proteins from the medium or the host cell.

### 30 C. Isolated MCP Proteins

- Another aspect of the invention pertains to isolated MCP proteins, and biologically active portions thereof. An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material when produced by recombinant DNA techniques, or chemical precursors or other chemicals when
- 35 chemically synthesized. The language "substantially free of cellular material" includes preparations of MCP protein in which the protein is separated from cellular components

language "substantially free of cellular material" includes preparations of MCP protein having less than about 30% (by dry weight) of non-MCP protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-MCP protein, still more preferably less than about 10% of non-MCP protein, and most preferably less than about 5% non-MCP protein. When the MCP protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation. The language "substantially free of chemical precursors or other chemicals" includes preparations of MCP protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of MCP protein having less than about 30% (by dry weight) of chemical precursors or non-MCP chemicals, more preferably less than about 20% chemical precursors or non-MCP chemicals, still more preferably less than about 10% chemical precursors or non-MCP chemicals, and most preferably less than about 5% chemical precursors or non-MCP chemicals. In preferred embodiments, isolated proteins or biologically active portions thereof lack contaminating proteins from the same organism from which the MCP protein is derived. Typically, such proteins are produced by recombinant expression of, for example, a *C. glutamicum* MCP protein in a microorganism such as *C. glutamicum*.

An isolated MCP protein or a portion thereof of the invention is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. In preferred embodiments, the protein or portion thereof comprises an amino acid sequence which is sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains the ability to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. The portion of the protein is preferably a biologically active portion as described herein. In another preferred embodiment, an MCP protein of the invention has an amino acid sequence shown in Appendix B. In yet another preferred embodiment, the MCP protein has an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, to a nucleotide sequence of Appendix A. In still another preferred embodiment, the MCP



protein has an amino acid sequence which is encoded by a nucleotide sequence that is at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, or 60%, preferably at least about 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%,  
5 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to one of the nucleic acid sequences of Appendix A, or a portion thereof. Ranges and identity values intermediate to the above-recited values, (e.g., 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For  
10 example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended to be included. The preferred MCP proteins of the present invention also preferably possess at least one of the MCP activities described herein. For example, a preferred MCP protein of the present invention includes an amino acid sequence encoded by a nucleotide sequence which  
15 hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence of Appendix A, and which is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms.

20 In other embodiments, the MCP protein is substantially homologous to an amino acid sequence of Appendix B and retains the functional activity of the protein of one of the sequences of Appendix B yet differs in amino acid sequence due to natural variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the MCP protein is a protein which comprises an amino acid sequence  
25 which is at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, or 60%, preferably at least about 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more  
30 homologous to an entire amino acid sequence of Appendix B and which has at least one of the MCP activities described herein. Ranges and identity values intermediate to the above-recited values, (e.g., 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended

Biologically active portions of an MCP protein include peptides comprising amino acid sequences derived from the amino acid sequence of an MCP protein, *e.g.*, an amino acid sequence shown in Appendix B or the amino acid sequence of a protein homologous to an MCP protein, which include fewer amino acids than a full length

5 MCP protein or the full length protein which is homologous to an MCP protein, and exhibit at least one activity of an MCP protein. Typically, biologically active portions (peptides, *e.g.*, peptides which are, for example, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) comprise a domain or motif with at least one activity of an MCP protein. Moreover, other biologically active portions, in which other  
10 regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the activities described herein. Preferably, the biologically active portions of an MCP protein include one or more selected domains/motifs or portions thereof having biological activity.

MCP proteins are preferably produced by recombinant DNA techniques. For  
15 example, a nucleic acid molecule encoding the protein is cloned into an expression vector (as described above), the expression vector is introduced into a host cell (as described above) and the MCP protein is expressed in the host cell. The MCP protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Alternative to recombinant expression, an MCP protein,  
20 polypeptide, or peptide can be synthesized chemically using standard peptide synthesis techniques. Moreover, native MCP protein can be isolated from cells (*e.g.*, endothelial cells, bacterial cells, fungal cells or other cells), for example using an anti-MCP antibody, which can be produced by standard techniques utilizing an MCP protein or fragment thereof of this invention.

25 The invention also provides MCP chimeric or fusion proteins. As used herein, an MCP "chimeric protein" or "fusion protein" comprises an MCP polypeptide operatively linked to a non-MCP polypeptide. An "MCP polypeptide" refers to a polypeptide having an amino acid sequence corresponding to an MCP protein, whereas a "non-MCP polypeptide" refers to a polypeptide having an amino acid sequence  
30 corresponding to a protein which is not substantially homologous to the MCP protein, *e.g.*, a protein which is different from the MCP protein and which is derived from the same or a different organism. Within the fusion protein, the term "operatively linked" is intended to indicate that the MCP polypeptide and the non-MCP polypeptide are fused in-frame to each other. The non-MCP polypeptide can be fused to the N-terminus or C-  
35 terminus of the MCP polypeptide. For example, in one embodiment the fusion protein is a GST-MCP fusion protein in which the MCP sequences are fused to the C-terminus

recombinant MCP proteins. In another embodiment, the fusion protein is an MCP protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells, bacterial host cells, fungal host cells), expression and/or secretion of an MCP protein can be increased through use of a heterologous signal sequence.

Preferably, an MCP chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). An MCP-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the MCP protein.

Homologues of the MCP protein can be generated by mutagenesis, e.g., discrete point mutation or truncation of the MCP protein. As used herein, the term "homologue" refers to a variant form of the MCP protein which acts as an agonist or antagonist of the activity of the MCP protein. An agonist of the MCP protein can retain substantially the same, or a subset, of the biological activities of the MCP protein. An antagonist of the MCP protein can inhibit one or more of the activities of the naturally occurring form of the MCP protein, by, for example, competitively binding to a downstream or upstream member of a biochemical pathway which includes the MCP protein.

In an alternative embodiment, homologues of the MCP protein can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the MCP protein for MCP protein agonist or antagonist activity. In one embodiment, a variegated library of MCP variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of MCP variants can be produced by, for example, enzymatically ligating a mixture of synthetic

fusion proteins (*e.g.*, for phage display) containing the set of MCP sequences therein. There are a variety of methods which can be used to produce libraries of potential MCP homologues from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the

5 synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential MCP sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science*  
10 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477.

In addition, libraries of fragments of the MCP protein coding can be used to generate a variegated population of MCP fragments for screening and subsequent selection of homologues of an MCP protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an  
15 MCP coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression  
20 vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the MCP protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for  
25 rapid screening of the gene libraries generated by the combinatorial mutagenesis of MCP homologues. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in  
30 which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify MCP homologues (Arkin and Yourvan (1992) *PNAS* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering*  
35 6(3):327-331).

*D. Uses and Methods of the Invention*

The nucleic acid molecules, proteins, protein homologues, fusion proteins, primers, vectors, and host cells described herein can be used in one or more of the following methods: identification of *C. glutamicum* and related organisms; mapping of genomes of organisms related to *C. glutamicum*; identification and localization of *C. glutamicum* sequences of interest; evolutionary studies; determination of MCP protein regions required for function; modulation of an MCP protein activity; modulation of the activity of one or more metabolic pathways; and modulation of cellular production of a desired compound, such as a fine chemical.

- 10       The MCP nucleic acid molecules of the invention have a variety of uses. First, they may be used to identify an organism as being *Corynebacterium glutamicum* or a close relative thereof. Also, they may be used to identify the presence of *C. glutamicum* or a relative thereof in a mixed population of microorganisms. The invention provides the nucleic acid sequences of a number of *C. glutamicum* genes, and probes based thereon; by probing the extracted genomic DNA of a culture of a unique or mixed population of microorganisms under stringent conditions with a probe spanning a region of a *C. glutamicum* gene which is unique to this organism, one can ascertain whether this organism is present. Although *Corynebacterium glutamicum* itself is nonpathogenic, it is related to pathogenic species, such as *Corynebacterium diphtheriae*. *Corynebacterium diphtheriae* is the causative agent of diphtheria, a rapidly developing, acute, febrile infection which involves both local and systemic pathology. In this disease, a local lesion develops in the upper respiratory tract and involves necrotic injury to epithelial cells; the bacilli secrete toxin which is disseminated through this lesion to distal susceptible tissues of the body. Degenerative changes brought about by the inhibition of protein synthesis in these tissues, which include heart, muscle, peripheral nerves, adrenals, kidneys, liver and spleen, result in the systemic pathology of the disease. Diphtheria continues to have high incidence in many parts of the world, including Africa, Asia, Eastern Europe and the independent states of the former Soviet Union. An ongoing epidemic of diphtheria in the latter two regions has resulted in at least 5,000 deaths since 1990.

- 35       In one embodiment, the invention provides a method of identifying the presence or activity of *Corynebacterium diphtheriae* in a subject. This method includes detection of one or more of the nucleic acid or amino acid sequences of the invention (e.g., the sequences set forth in Appendix A or Appendix B) in a subject, thereby detecting the presence or activity of *Corynebacterium diphtheriae* in the subject. *C. glutamicum* and *C. diphtheriae* are related bacteria, and many of the nucleic acid and protein molecules

in *C. glutamicum* are homologous to *C. diphtheriae* nucleic acid and protein molecules, and can therefore be used to detect *C. diphtheriae* in a subject.

To detect the presence of *C. glutamicum* in a sample, techniques well known in the art may be employed. Specifically, the cells in the sample may optionally first be

5 cultured in a suitable liquid or on a suitable solid culture medium to increase the number of cells in the sample. These cells are lysed, and the total DNA content extracted and optionally purified to remove debris and protein material which may interfere with subsequent analysis. The polymerase chain reaction or a similar technique known in the art is performed (for general reference on methodologies commonly used for the

10 amplification of nucleic acid sequences, see Mullis *et al.*, U.S. Patent No. 4,683,195, Mullis *et al.*, U.S. Patent No. 4,965,188, and Innis, M.A., and Gelfand, D. H., (1989) PCR Protocols, A guide to Methods and Applications, Academic Press, p. 3-12, and (1988) Biotechnology 6:1197, and International Patent Application No. WO89/01050) in which primers specific to an MCP nucleic acid molecule of the invention are

15 incubated with the nucleic acid sample such that, if present in the sample, that particular MCP nucleic acid sequence will be amplified. The particular MCP nucleic acid to be amplified is selected based on its uniqueness to the *C. glutamicum* genome, or to the genomes of *C. glutamicum* and only a few closely related bacteria. The presence of the desired amplified product is thus indicative of the presence of *C. glutamicum*, or an

20 organism closely related to *C. glutamicum*.

Further, the nucleic acid and protein molecules of the invention may serve as markers for specific regions of the genome. It is possible, using techniques well known in the art, to ascertain the physical location on the *C. glutamicum* genome of the MCP nucleic acid molecules of the invention, which in turn provides markers on the genome

25 which can be used to aid in the placement of other nucleic acid molecules and genes on the genome map. Also, the nucleic acid molecules of the invention may be sufficiently homologous to the sequences of related bacterial species that these nucleic acid molecules may similarly permit the construction of a genomic map in such bacteria (*e.g.*, *Brevibacterium lactofermentum*).

30 The nucleic acid molecules of the invention have utility not only in the mapping of the genome, but also for functional studies of *C. glutamicum* proteins. For example, to identify the region of the genome to which a particular *C. glutamicum* DNA-binding protein binds, the *C. glutamicum* genome could be digested, and the fragments incubated with the DNA-binding protein. Those which bind the protein may be additionally probed

35 with the nucleic acid molecules of the invention, preferably with readily detectable labels; binding of such a nucleic acid molecule to the genome fragment enables the

multiple times with different enzymes, facilitates a rapid determination of the nucleic acid sequence to which the protein binds.

The MCP nucleic acid molecules of the invention are also useful for evolutionary and protein structural studies. The metabolic processes in which the molecules of the invention participate are utilized by a wide variety of prokaryotic and eukaryotic cells; by comparing the sequences of the nucleic acid molecules of the present invention to those encoding similar enzymes from other organisms, the evolutionary relatedness of the organisms can be assessed. Similarly, such a comparison permits an assessment of which regions of the sequence are conserved and which are not, which may aid in determining those regions of the protein which are essential for the functioning of the enzyme. This type of determination is of value for protein engineering studies and may give an indication of what the protein can tolerate in terms of mutagenesis without losing function.

The MCP protein molecules of the invention may also be utilized as markers for the classification of an unknown bacterium as *C. glutamicum*, or for the identification of *C. glutamicum* or closely related bacteria in a sample. For example, using techniques well known in the art, cells in a sample may optionally be amplified (*e.g.*, by culturing in an appropriate medium) to increase the sample size, and then may be lysed to release proteins contained therein. This sample may optionally be purified to remove debris and nucleic acid molecules which may interfere with subsequent analysis. Antibodies specific for a selected MCP protein of the invention may be incubated with the protein sample in a typical Western assay format (see, *e.g.*, Ausubel *et al.*, (1988) Current Protocols in Molecular Biology, Wiley: New York) in which the antibody will bind to its target protein if this protein is present in the sample. An MCP protein is selected for this type of assay if it is unique or nearly unique to *C. glutamicum* or *C. glutamicum* and bacteria very closely related to *C. glutamicum*. Proteins in the sample are then separated by gel electrophoresis, and transferred to a suitable matrix, such as nitrocellulose. An appropriate secondary antibody having a detectable label (*e.g.*, chemiluminescent or colorimetric) is incubated with this matrix, followed by stringent washing. The presence or absence of the label is indicative of the presence or absence of the target protein in the sample. If the protein is present, then this is indicative of the presence of *C. glutamicum*. A similar process enables the classification of an unknown bacterium as *C. glutamicum*; if a panel of proteins specific to *C. glutamicum* are not detected in protein samples prepared from the unknown bacterium, then that bacterium is not likely to be *C. glutamicum*.

The invention provides methods for screening molecules which modulate the activity of an MCP protein, either by interacting with the protein itself or a substrate or

binding partner of the MCP protein, or by modulating the transcription or translation of an MCP nucleic acid molecule of the invention. In such methods, a microorganism expressing one or more MCP proteins of the invention is contacted with one or more test compounds, and the effect of each test compound on the activity or level of expression

5 of the MCP protein is assessed.

Genetic manipulation of the MCP nucleic acid molecules of the invention may result in the production of MCP proteins having functional differences from the wild-type MCP proteins. These proteins may be improved in efficiency or activity, may be present in greater numbers in the cell than is usual, or may be decreased in efficiency or  
10 activity.

Such changes in activity may directly modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*. For example, by modifying the activity of a protein involved in the biosynthesis or degradation of a fine chemical (*i.e.*, through mutagenesis of the corresponding gene),  
15 one may directly modulate the ability of the cell to synthesize or to degrade this compound, thereby modulating the yield and/or efficiency of production of the fine chemical. Similarly, by modulating the activity of a protein which regulates a fine chemical metabolic pathway, one may directly influence whether the production of the desired compound is up- or down-regulated, either of which will modulate the yield or  
20 efficiency of production of the fine chemical from the cell.

Indirect modulation of fine chemical production may also result by modifying the activity of a protein of the invention (*i.e.*, by mutagenesis of the corresponding gene) such that the overall ability of the cell to grow and divide or to remain viable and productive is increased. The production of fine chemicals from *C. glutamicum* is  
25 generally accomplished by the large-scale fermentative culture of these microorganisms, conditions which are frequently suboptimal for growth and cell division. By engineering a protein of the invention (*e.g.*, a stress response protein, a cell wall protein, or proteins involved in the metabolism of compounds necessary for cell growth and division to occur, such as nucleotides and amino acids) such that it is better able to  
30 survive, grow, and multiply in such conditions, it may be possible to increase the number and productivity of such engineered *C. glutamicum* cells in large-scale culture, which in turn should result in increased yields and/or efficiency of production of one or more desired fine chemicals. Further, the metabolic pathways of any cell are necessarily interrelated and coregulated. By altering the activity or regulation of any one metabolic  
35 pathway in *C. glutamicum* (*i.e.*, by altering the activity of one of the proteins of the invention which participates in such a pathway), it is possible to concomitantly alter the



activity or regulation of other metabolic pathways in this microorganism, which may be directly involved in the synthesis or degradation of a fine chemical.

The aforementioned mutagenesis strategies for MCP proteins to result in increased yields of a fine chemical from *C. glutamicum* are not meant to be limiting;

5 variations on these strategies will be readily apparent to one of ordinary skill in the art. Using such strategies, and incorporating the mechanisms disclosed herein, the nucleic acid and protein molecules of the invention may be utilized to generate *C. glutamicum* or related strains of bacteria expressing mutated MCP nucleic acid and protein molecules such that the yield, production, and/or efficiency of production of a desired  
10 compound is improved. This desired compound may be any natural product of *C. glutamicum*, which includes the final products of biosynthesis pathways and intermediates of naturally-occurring metabolic pathways, as well as molecules which do not naturally occur in the metabolism of *C. glutamicum*, but which are produced by a *C. glutamicum* strain of the invention.

15 This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patent applications, patents, published patent applications, Tables, Appendices, and the sequence listing cited throughout this application are hereby incorporated by reference.

## 20 Exemplification

### **Example 1: Preparation of total genomic DNA of *Corynebacterium glutamicum* ATCC 13032**

25 A culture of *Corynebacterium glutamicum* (ATCC 13032) was grown overnight at 30°C with vigorous shaking in BHI medium (Difco). The cells were harvested by centrifugation, the supernatant was discarded and the cells were resuspended in 5 ml buffer-I (5% of the original volume of the culture — all indicated volumes have been calculated for 100 ml of culture volume). Composition of buffer-I: 140.34 g/l sucrose,  
30 2.46 g/l  $\text{MgSO}_4 \times 7\text{H}_2\text{O}$ , 10 ml/l  $\text{KH}_2\text{PO}_4$  solution (100 g/l, adjusted to pH 6.7 with KOH), 50 ml/l M12 concentrate (10 g/l  $(\text{NH}_4)_2\text{SO}_4$ , 1 g/l NaCl, 2 g/l  $\text{MgSO}_4 \times 7\text{H}_2\text{O}$ , 0.2 g/l  $\text{CaCl}_2$ , 0.5 g/l yeast extract (Difco), 10 ml/l trace-elements-mix (200 mg/l  $\text{FeSO}_4 \times \text{H}_2\text{O}$ , 10 mg/l  $\text{ZnSO}_4 \times 7 \text{H}_2\text{O}$ , 3 mg/l  $\text{MnCl}_2 \times 4 \text{H}_2\text{O}$ , 30 mg/l  $\text{H}_3\text{BO}_3$ , 20 mg/l  $\text{CoCl}_2 \times 6 \text{H}_2\text{O}$ , 1 mg/l  $\text{NiCl}_2 \times 6 \text{H}_2\text{O}$ , 3 mg/l  $\text{Na}_2\text{MoO}_4 \times 2 \text{H}_2\text{O}$ , 500 mg/l complexing agent  
35 (EDTA or critic acid), 100 ml/l vitamins-mix (0.2 mg/l biotin, 0.2 mg/l folic acid, 20 mg/l p-amino benzoic acid, 20 mg/l riboflavin, 40 mg/l ca-panthothenate, 140 mg/l nicotinic acid, 40 mg/l pyridoxole hydrochloride, 200 mg/l myo-inositol). Lysozyme

was added to the suspension to a final concentration of 2.5 mg/ml. After an approximately 4 h incubation at 37°C, the cell wall was degraded and the resulting protoplasts are harvested by centrifugation. The pellet was washed once with 5 ml buffer-I and once with 5-ml TE-buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). The

5 pellet was resuspended in 4 ml TE-buffer and 0.5 ml SDS solution (10%) and 0.5 ml NaCl solution (5 M) are added. After adding of proteinase K to a final concentration of 200 µg/ml, the suspension is incubated for ca. 18 h at 37°C. The DNA was purified by extraction with phenol, phenol-chloroform-isoamylalcohol and chloroform-isoamylalcohol using standard procedures. Then, the DNA was precipitated by adding  
10 1/50 volume of 3 M sodium acetate and 2 volumes of ethanol, followed by a 30 min incubation at -20°C and a 30 min centrifugation at 12,000 rpm in a high speed centrifuge using a SS34 rotor (Sorvall). The DNA was dissolved in 1 ml TE-buffer containing 20 µg/ml RNaseA and dialysed at 4°C against 1000 ml TE-buffer for at least 3 hours. During this time, the buffer was exchanged 3 times. To aliquots of 0.4 ml of the  
15 dialysed DNA solution, 0.4 ml of 2 M LiCl and 0.8 ml of ethanol are added. After a 30 min incubation at -20°C, the DNA was collected by centrifugation (13,000 rpm, Biofuge Fresco, Heraeus, Hanau, Germany). The DNA pellet was dissolved in TE-buffer. DNA prepared by this procedure could be used for all purposes, including southern blotting or construction of genomic libraries.

20

**Example 2: Construction of genomic libraries in *Escherichia coli* of *Corynebacterium glutamicum* ATCC13032.**

Using DNA prepared as described in Example 1, cosmid and plasmid libraries were constructed according to known and well established methods (see e.g., Sambrook, J. *et al.*  
25 (1989) "Molecular Cloning : A Laboratory Manual", Cold Spring Harbor Laboratory Press, or Ausubel, F.M. *et al.* (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.)

Any plasmid or cosmid could be used. Of particular use were the plasmids pBR322 (Sutcliffe, J.G. (1979) *Proc. Natl. Acad. Sci. USA*, 75:3737-3741); pACYC177 (Change &  
30 Cohen (1978) *J. Bacteriol* 134:1141-1156), plasmids of the pBS series (pBSSK+, pBSSK- and others; Stratagene, LaJolla, USA), or cosmids as SuperCos1 (Stratagene, LaJolla, USA) or Lorist6 (Gibson, T.J., Rosenthal A. and Waterson, R.H. (1987) *Gene* 53:283-286. Gene libraries specifically for use in *C. glutamicum* may be constructed using plasmid pSL109 (Lee, H.-S. and

**Example 3: DNA Sequencing and Computational Functional Analysis**

*Sub B1* Genomic libraries as described in Example 2 were used for DNA sequencing according to standard methods, in particular by the chain termination method using

5 ABI377 sequencing machines (see e.g., Fleischman, R.D. *et al.* (1995) "Whole-genome Random Sequencing and Assembly of Haemophilus Influenzae Rd., *Science*, 269:496-512). Sequencing primers with the following nucleotide sequences were used: 5'-GGAAACAGTATGACCATG-3' or 5'-GTAAAACGACGGCCAGT-3'.

**10 Example 4: *In vivo* Mutagenesis**

*In vivo* mutagenesis of *Corynebacterium glutamicum* can be performed by passage of plasmid (or other vector) DNA through *E. coli* or other microorganisms (e.g. *Bacillus* spp. or yeasts such as *Saccharomyces cerevisiae*) which are impaired in their capabilities to maintain the integrity of their genetic information. Typical mutator strains have mutations in the genes

15 for the DNA repair system (e.g., mutHLS, mutD, mutT, etc.; for reference, see Rupp, W.D. (1996) DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, p. 2277-2294, ASM: Washington.) Such strains are well known to those of ordinary skill in the art. The use of such strains is illustrated, for example, in Greener, A. and Callahan, M. (1994) *Strategies* 7: 32-34.

**20 Example 5: DNA Transfer Between *Escherichia coli* and *Corynebacterium glutamicum***

Several *Corynebacterium* and *Brevibacterium* species contain endogenous plasmids (as e.g., pHM1519 or pBL1) which replicate autonomously (for review see, e.g., Martin, J.F. *et al.* (1987) *Biotechnology*, 5:137-146). Shuttle vectors for *Escherichia coli*

25 and *Corynebacterium glutamicum* can be readily constructed by using standard vectors for *E. coli* (Sambrook, J. *et al.* (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press or Ausubel, F.M. *et al.* (1994) "Current Protocols in Molecular Biology", John Wiley & Sons) to which a origin or replication for and a suitable marker from *Corynebacterium glutamicum* is added. Such origins of replication

30 are preferably taken from endogenous plasmids isolated from *Corynebacterium* and *Brevibacterium* species. Of particular use as transformation markers for these species are genes for kanamycin resistance (such as those derived from the Tn5 or Tn903 transposons) or chloramphenicol (Winnacker, E.L. (1987) "From Genes to Clones — Introduction to Gene Technology, VCH, Weinheim). There are numerous examples in the

35 literature of the construction of a wide variety of shuttle vectors which replicate in both *E. coli* and *C. glutamicum*, and which can be used for several purposes, including gene over-expression (for reference, see e.g., Yoshihama, M. *et al.* (1985) *J. Bacteriol.* 162:591-597,

*Martin J.F. et al. (1987) Biotechnology, 5:137-146 and Eikmanns, B.J. et al. (1991) Gene, 102:93-98).*

Using standard methods, it is possible to clone a gene of interest into one of the shuttle-vectors described above and to introduce such a hybrid vectors into strains of

5 *Corynebacterium glutamicum*. Transformation of *C. glutamicum* can be achieved by protoplast transformation (Kastsumata, R. *et al.* (1984) *J. Bacteriol.* 159:306-311), electroporation (Liebl, E. *et al.* (1989) *FEMS Microbiol. Letters*, 53:399-303) and in cases where special vectors are used, also by conjugation (as described *e.g.* in Schäfer, A *et al.* (1990) *J. Bacteriol.* 172:1663-1666). It is also possible to transfer the shuttle vectors for  
10 *C. glutamicum* to *E. coli* by preparing plasmid DNA from *C. glutamicum* (using standard methods well-known in the art) and transforming it into *E. coli*. This transformation step can be performed using standard methods, but it is advantageous to use an *Mcr*-deficient *E. coli* strain, such as NM522 (Gough & Murray (1983) *J. Mol. Biol.* 166:1-19).

Genes may be overexpressed in *C. glutamicum* strains using plasmids which  
15 comprise pCG1 (U.S. Patent No. 4,617,267) or fragments thereof, and optionally the gene for kanamycin resistance from TN903 (Grindley, N.D. and Joyce, C.M. (1980) *Proc. Natl. Acad. Sci. USA* 77(12): 7176-7180). In addition, genes may be overexpressed in *C. glutamicum* strains using plasmid pSL109 (Lee, H.-S. and A. J. Sinskey (1994) *J. Microbiol. Biotechnol.* 4: 256-263).

20 Aside from the use of replicative plasmids, gene overexpression can also be achieved by integration into the genome. Genomic integration in *C. glutamicum* or other *Corynebacterium* or *Brevibacterium* species may be accomplished by well-known methods, such as homologous recombination with genomic region(s), restriction endonuclease mediated integration (REMI) (see, *e.g.*, DE Patent 19823834), or through  
25 the use of transposons. It is also possible to modulate the activity of a gene of interest by modifying the regulatory regions (*e.g.*, a promoter, a repressor, and/or an enhancer) by sequence modification, insertion, or deletion using site-directed methods (such as homologous recombination) or methods based on random events (such as transposon mutagenesis or REMI). Nucleic acid sequences which function as transcriptional  
30 terminators may also be inserted 3' to the coding region of one or more genes of the invention; such terminators are well-known in the art and are described, for example, in Winnacker, E.L. (1987) *From Genes to Clones – Introduction to Gene Technology*. VCH: Weinheim.

### 35 **Example 6: Assessment of the Expression of the Mutant Protein**

Observations of the activity of a mutated protein in a transformed host cell rely on the fact that the mutant protein is expressed in a similar fashion and in a similar quantity

to that of the wild-type protein. A useful method to ascertain the level of transcription of the mutant gene (an indicator of the amount of mRNA available for translation to the gene product) is to perform a Northern blot (for reference see, for example, Ausubel *et al.*

(1988) *Current Protocols in Molecular Biology*, Wiley: New York), in which a primer

- 5 designed to bind to the gene of interest is labeled with a detectable tag (usually radioactive or chemiluminescent), such that when the total RNA of a culture of the organism is extracted, run on gel, transferred to a stable matrix and incubated with this probe, the binding and quantity of binding of the probe indicates the presence and also the quantity of mRNA for this gene. This information is evidence of the degree of transcription of the
- 10 mutant gene. Total cellular RNA can be prepared from *Corynebacterium glutamicum* by several methods, all well-known in the art, such as that described in Bormann, E.R. *et al.* (1992) *Mol. Microbiol.* 6: 317-326.

- To assess the presence or relative quantity of protein translated from this mRNA, standard techniques, such as a Western blot, may be employed (see, for example, Ausubel
- 15 *et al.* (1988) *Current Protocols in Molecular Biology*, Wiley: New York). In this process, total cellular proteins are extracted, separated by gel electrophoresis, transferred to a matrix such as nitrocellulose, and incubated with a probe, such as an antibody, which specifically binds to the desired protein. This probe is generally tagged with a chemiluminescent or colorimetric label which may be readily detected. The presence and
- 20 quantity of label observed indicates the presence and quantity of the desired mutant protein present in the cell.

#### 25 **Example 7: Growth of Genetically Modified *Corynebacterium glutamicum* — Media and Culture Conditions**

- Genetically modified *Corynebacteria* are cultured in synthetic or natural growth media. A number of different growth media for *Corynebacteria* are both well-known and readily available (Lieb *et al.* (1989) *Appl. Microbiol. Biotechnol.*, 32:205-210; von der
- 30 Osten *et al.* (1998) *Biotechnology Letters*, 11:11-16; Patent DE 4,120,867; Liebl (1992) "The Genus *Corynebacterium*, in: *The Procaryotes*, Volume II, Balows, A. *et al.*, eds. Springer-Verlag). These media consist of one or more carbon sources, nitrogen sources, inorganic salts, vitamins and trace elements. Preferred carbon sources are sugars, such as mono-, di-, or polysaccharides. For example, glucose, fructose, mannose, galactose, ribose, sorbose, ribulose, lactose, maltose, sucrose, raffinose, starch or cellulose serve as
- 35 very good carbon sources. It is also possible to supply sugar to the media via complex compounds such as molasses or other by-products from sugar refinement. It can also be advantageous to supply mixtures of different carbon sources. Other possible carbon

sources are alcohols and organic acids, such as methanol, ethanol, acetic acid or lactic acid. Nitrogen sources are usually organic or inorganic nitrogen compounds, or materials which contain these compounds. Exemplary nitrogen sources include ammonia gas or ammonia salts, such as  $\text{NH}_4\text{Cl}$  or  $(\text{NH}_4)_2\text{SO}_4$ ,  $\text{NH}_4\text{OH}$ , nitrates, urea, amino acids or complex nitrogen sources like corn steep liquor, soy bean flour, soy bean protein, yeast extract, meat extract and others.

Inorganic salt compounds which may be included in the media include the chloride-, phosphorous- or sulfate- salts of calcium, magnesium, sodium, cobalt, molybdenum, potassium, manganese, zinc, copper and iron. Chelating compounds can be added to the medium to keep the metal ions in solution. Particularly useful chelating compounds include dihydroxyphenols, like catechol or protocatechuate, or organic acids, such as citric acid. It is typical for the media to also contain other growth factors, such as vitamins or growth promoters, examples of which include biotin, riboflavin, thiamin, folic acid, nicotinic acid, pantothenate and pyridoxin. Growth factors and salts frequently originate from complex media components such as yeast extract, molasses, corn steep liquor and others. The exact composition of the media compounds depends strongly on the immediate experiment and is individually decided for each specific case. Information about media optimization is available in the textbook "Applied Microbiol. Physiology, A Practical Approach" (eds. P.M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0 19 963577 3). It is also possible to select growth media from commercial suppliers, like standard 1 (Merck) or BHI (grain heart infusion, DIFCO) or others.

All medium components are sterilized, either by heat (20 minutes at 1.5 bar and  $121^\circ\text{C}$ ) or by sterile filtration. The components can either be sterilized together or, if necessary, separately. All media components can be present at the beginning of growth, or they can optionally be added continuously or batchwise.

Culture conditions are defined separately for each experiment. The temperature should be in a range between  $15^\circ\text{C}$  and  $45^\circ\text{C}$ . The temperature can be kept constant or can be altered during the experiment. The pH of the medium should be in the range of 5 to 8.5, preferably around 7.0, and can be maintained by the addition of buffers to the media. An exemplary buffer for this purpose is a potassium phosphate buffer. Synthetic buffers such as MOPS, HEPES, ACES and others can alternatively or simultaneously be used. It is also possible to maintain a constant culture pH through the addition of  $\text{NaOH}$  or  $\text{NH}_4\text{OH}$  during growth. If complex medium components such as yeast extract are utilized, the necessity for additional buffers may be reduced, due to the fact that many complex compounds have high buffer capacities. If a fermentor is utilized for culturing the micro-

The incubation time is usually in a range from several hours to several days. This time is selected in order to permit the maximal amount of product to accumulate in the broth. The disclosed growth experiments can be carried out in a variety of vessels, such as microtiter plates, glass tubes, glass flasks or glass or metal fermentors of different sizes.

- 5 For screening a large number of clones, the microorganisms should be cultured in microtiter plates, glass tubes or shake flasks, either with or without baffles. Preferably 100 ml shake flasks are used, filled with 10% (by volume) of the required growth medium. The flasks should be shaken on a rotary shaker (amplitude 25 mm) using a speed-range of 100 – 300 rpm. Evaporation losses can be diminished by the maintenance  
10 of a humid atmosphere; alternatively, a mathematical correction for evaporation losses should be performed.

- If genetically modified clones are tested, an unmodified control clone or a control clone containing the basic plasmid without any insert should also be tested. The medium is inoculated to an OD<sub>600</sub> of 0.5 – 1.5 using cells grown on agar plates, such as CM plates  
15 (10 g/l glucose, 2.5 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract, 22 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract, 22 g/l agar, pH 6.8 with 2M NaOH) that had been incubated at 30°C. Inoculation of the media is accomplished by either introduction of a saline suspension of *C. glutamicum* cells from CM plates or addition of a liquid preculture of this bacterium.

20

### **Example 8 – *In vitro* Analysis of the Function of Mutant Proteins**

- The determination of activities and kinetic parameters of enzymes is well established in the art. Experiments to determine the activity of any given altered enzyme must be tailored to the specific activity of the wild-type enzyme, which is well  
25 within the ability of one of ordinary skill in the art. Overviews about enzymes in general, as well as specific details concerning structure, kinetics, principles, methods, applications and examples for the determination of many enzyme activities may be found, for example, in the following references: Dixon, M., and Webb, E.C., (1979) *Enzymes*. Longmans: London; Fersht, (1985) *Enzyme Structure and Mechanism*. Freeman: New York; Walsh, (1979) *Enzymatic Reaction Mechanisms*. Freeman: San  
30 Francisco; Price, N.C., Stevens, L. (1982) *Fundamentals of Enzymology*. Oxford Univ. Press: Oxford; Boyer, P.D., ed. (1983) *The Enzymes*, 3<sup>rd</sup> ed. Academic Press: New York; Bisswanger, H., (1994) *Enzymkinetik*, 2<sup>nd</sup> ed. VCH: Weinheim (ISBN 3527300325); Bergmeyer, H.U., Bergmeyer, J., Graßl, M., eds. (1983-1986) *Methods of Enzymatic Analysis*, 3<sup>rd</sup> ed., vol. I-XII, Verlag Chemie: Weinheim; and Ullmann's

The activity of proteins which bind to DNA can be measured by several well-established methods, such as DNA band-shift assays (also called gel retardation assays). The effect of such proteins on the expression of other molecules can be measured using reporter gene assays (such as that described in Kolmar, H. *et al.* (1995) *EMBO J.* 14:

- 5 3895-3904 and references cited therein). Reporter gene test systems are well known and established for applications in both pro- and eukaryotic cells, using enzymes such as beta-galactosidase, green fluorescent protein, and several others.

- The determination of activity of membrane-transport proteins can be performed according to techniques such as those described in Gennis, R.B. (1989) "Pores,  
10 Channels and Transporters", in *Biomembranes, Molecular Structure and Function*, Springer: Heidelberg, p. 85-137; 199-234; and 270-322.

**Example 9: Analysis of Impact of Mutant Protein on the Production of the Desired Product**

- 15 The effect of the genetic modification in *C. glutamicum* on production of a desired compound (such as an amino acid) can be assessed by growing the modified microorganism under suitable conditions (such as those described above) and analyzing the medium and/or the cellular component for increased production of the desired product (*i.e.*, an amino acid). Such analysis techniques are well known to one of  
20 ordinary skill in the art, and include spectroscopy, thin layer chromatography, staining methods of various kinds, enzymatic and microbiological methods, and analytical chromatography such as high performance liquid chromatography (see, for example, Ullman, *Encyclopedia of Industrial Chemistry*, vol. A2, p. 89-90 and p. 443-613, VCH: Weinheim (1985); Fallon, A. *et al.*, (1987) "Applications of HPLC in Biochemistry" in:  
25 Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17; Rehm *et al.* (1993) *Biotechnology*, vol. 3, Chapter III: "Product recovery and purification", page 469-714, VCH: Weinheim; Belter, P.A. *et al.* (1988) *Bioseparations: downstream processing for biotechnology*, John Wiley and Sons; Kennedy, J.F. and Cabral, J.M.S. (1992) *Recovery processes for biological materials*, John Wiley and Sons; Shaeiwitz,  
30 J.A. and Henry, J.D. (1988) *Biochemical separations*, in: *Ullmann's Encyclopedia of Industrial Chemistry*, vol. B3, Chapter 11, page 1-27, VCH: Weinheim; and Dechow, F.J. (1989) *Separation and purification techniques in biotechnology*, Noyes Publications.)

- In addition to the measurement of the final product of fermentation, it is also  
35 possible to analyze other components of the metabolic pathways utilized for the production of the desired compound, such as intermediates and side-products, to determine the overall efficiency of production of the compound. Analysis methods



include measurements of nutrient levels in the medium (*e.g.*, sugars, hydrocarbons, nitrogen sources, phosphate, and other ions), measurements of biomass composition and growth, analysis of the production of common metabolites of biosynthetic pathways, and measurement of gasses produced during fermentation. Standard methods for these

- 5 measurements are outlined in Applied Microbial Physiology, A Practical Approach, P.M. Rhodes and P.F. Stanbury, eds., IRL Press, p. 103-129; 131-163; and 165-192 (ISBN: 0199635773) and references cited therein.

**Example 10: Purification of the Desired Product from *C. glutamicum* Culture**

- 10 Recovery of the desired product from the *C. glutamicum* cells or supernatant of the above-described culture can be performed by various methods well known in the art. If the desired product is not secreted from the cells, the cells can be harvested from the culture by low-speed centrifugation, the cells can be lysed by standard techniques, such as mechanical force or sonication. The cellular debris is removed by centrifugation, and
- 15 the supernatant fraction containing the soluble proteins is retained for further purification of the desired compound. If the product is secreted from the *C. glutamicum* cells, then the cells are removed from the culture by low-speed centrifugation, and the supernate fraction is retained for further purification.

- The supernatant fraction from either purification method is subjected to
- 20 chromatography with a suitable resin, in which the desired molecule is either retained on a chromatography resin while many of the impurities in the sample are not, or where the impurities are retained by the resin while the sample is not. Such chromatography steps may be repeated as necessary, using the same or different chromatography resins. One of ordinary skill in the art would be well-versed in the selection of appropriate
- 25 chromatography resins and in their most efficacious application for a particular molecule to be purified. The purified product may be concentrated by filtration or ultrafiltration, and stored at a temperature at which the stability of the product is maximized.

- There are a wide array of purification methods known to the art and the preceding method of purification is not meant to be limiting. Such purification
- 30 techniques are described, for example, in Bailey, J.E. & Ollis, D.F. Biochemical Engineering Fundamentals, McGraw-Hill: New York (1986).

- The identity and purity of the isolated compounds may be assessed by techniques standard in the art. These include high-performance liquid chromatography (HPLC), spectroscopic methods, staining methods, thin layer chromatography, NIRS, enzymatic
- 35 assay, or microbiologically. Such analysis methods are reviewed in: Patek *et al.* (1994) *Appl. Environ. Microbiol.* 60: 133-140; Malakhova *et al.* (1996) *Biotekhnologiya* 11: 27-10, 67-70; Ulmann's Encyclopedia

of Industrial Chemistry, (1996) vol. A27, VCH: Weinheim, p. 89-90, p. 521-540, p. 540-547, p. 559-566, 575-581 and p. 581-587; Michal, G. (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A. *et al.* (1987) *Applications of HPLC in Biochemistry in: Laboratory Techniques in*

5 *Biochemistry and Molecular Biology*, vol. 17.

### Example 11: Analysis of the Gene Sequences of the Invention

The comparison of sequences and determination of percent homology between two sequences are art-known techniques, and can be accomplished using a mathematical algorithm, such as the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-68, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-77. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to MCP nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to MCP protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, one of ordinary skill in the art will know how to optimize the parameters of the program (*e.g.*, XBLAST and NBLAST) for the specific sequence being analyzed.

Another example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Meyers and Miller ((1988) *Comput. Appl. Biosci.* 4: 11-17). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Additional algorithms for sequence analysis are known in the art, and include ADVANCE and ADAM. described in Torelli and Robotti (1994) *Comput. Appl. Biosci.* 10:3-5; and FASTA, described in Pearson and Lipman (1988) *P.N.A.S.* 85:2444-8.

35 *SUB P1* The percent homology between two amino acid sequences can also be accomplished using the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 12, 10, 8, 6, or 4 and a length weight of 2, 3, or 4. The percent homology between two nucleic acid sequences can be accomplished using the GAP program in the

GCG software package, using standard parameters, such as a gap weight of 50 and a length weight of 3.

A comparative analysis of the gene sequences of the invention with those present in Genbank has been performed using techniques known in the art (see, *e.g.*, Bexevanis and Ouellette, eds. (1998) *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. John Wiley and Sons: New York). The gene sequences of the invention were compared to genes present in Genbank in a three-step process. In a first step, a BLASTN analysis (*e.g.*, a local alignment analysis) was performed for each of the sequences of the invention against the nucleotide sequences present in Genbank, and the top 500 hits were retained for further analysis. A subsequent FASTA search (*e.g.*, a combined local and global alignment analysis, in which limited regions of the sequences are aligned) was performed on these 500 hits. Each gene sequence of the invention was subsequently globally aligned to each of the top three FASTA hits, using the GAP program in the GCG software package (using standard parameters). In order to obtain correct results, the length of the sequences extracted from Genbank were adjusted to the length of the query sequences by methods well-known in the art. The results of this analysis are set forth in Table 4. The resulting data is identical to that which would have been obtained had a GAP (global) analysis alone been performed on each of the genes of the invention in comparison with each of the references in Genbank, but required significantly reduced computational time as compared to such a database-wide GAP (global) analysis. Sequences of the invention for which no alignments above the cutoff values were obtained are indicated on Table 4 by the absence of alignment information. It will further be understood by one of ordinary skill in the art that the GAP alignment homology percentages set forth in Table 4 under the heading "% homology (GAP)" are listed in the European numerical format, wherein a ',' represents a decimal point. For example, a value of "40,345" in this column represents "40.345%".

#### **Example 12: Construction and Operation of DNA Microarrays**

The sequences of the invention may additionally be used in the construction and application of DNA microarrays (the design, methodology, and uses of DNA arrays are well known in the art, and are described, for example, in Schena, M. *et al.* (1995) *Science* 270: 467-470; Wodicka, L. *et al.* (1997) *Nature Biotechnology* 15: 1359-1367; DeSaizieu, A. *et al.* (1998) *Nature Biotechnology* 16: 45-48; and DeRisi, J.L. *et al.* (1997) *Science* 278: 680-686).

DNA microarrays are solid or flexible supports consisting of nitrocellulose, nylon, glass, silicone, or other materials. Nucleic acid molecules may be attached to the

acid mixtures can be hybridized to the immobilized nucleic acid molecules, and the label may be used to monitor and measure the individual signal intensities of the hybridized molecules at defined regions. This methodology allows the simultaneous quantification of the relative or absolute amount of all or selected nucleic acids in the applied nucleic

- 5 acid sample or mixture. DNA microarrays, therefore, permit an analysis of the expression of multiple (as many as 6800 or more) nucleic acids in parallel (see, *e.g.*, Schena, M. (1996) *BioEssays* 18(5): 427-431).

The sequences of the invention may be used to design oligonucleotide primers which are able to amplify defined regions of one or more *C. glutamicum* genes by a  
10 nucleic acid amplification reaction such as the polymerase chain reaction. The choice and design of the 5' or 3' oligonucleotide primers or of appropriate linkers allows the covalent attachment of the resulting PCR products to the surface of a support medium described above (and also described, for example, Schena, M. *et al.* (1995) *Science* 270: 467-470).

- 15 Nucleic acid microarrays may also be constructed by *in situ* oligonucleotide synthesis as described by Wodicka, L. *et al.* (1997) *Nature Biotechnology* 15: 1359-1367. By photolithographic methods, precisely defined regions of the matrix are exposed to light. Protective groups which are photolabile are thereby activated and undergo nucleotide addition, whereas regions that are masked from light do not undergo  
20 any modification. Subsequent cycles of protection and light activation permit the synthesis of different oligonucleotides at defined positions. Small, defined regions of the genes of the invention may be synthesized on microarrays by solid phase oligonucleotide synthesis.

The nucleic acid molecules of the invention present in a sample or mixture of  
25 nucleotides may be hybridized to the microarrays. These nucleic acid molecules can be labeled according to standard methods. In brief, nucleic acid molecules (*e.g.*, mRNA molecules or DNA molecules) are labeled by the incorporation of isotopically or fluorescently labeled nucleotides, *e.g.*, during reverse transcription or DNA synthesis. Hybridization of labeled nucleic acids to microarrays is described (*e.g.*, in Schena, M. *et al.* (1995) *supra*; Wodicka, L. *et al.* (1997), *supra*; and DeSaizieu A. *et al.* (1998),  
30 *supra*). The detection and quantification of the hybridized molecule are tailored to the specific incorporated label. Radioactive labels can be detected, for example, as described in Schena, M. *et al.* (1995) *supra*) and fluorescent labels may be detected, for example, by the method of Shalon *et al.* (1996) *Genome Research* 6: 639-645).

- 35 The application of the sequences of the invention to DNA microarray technology, as described above, permits comparative analyses of different strains of *C.*

based on individual transcript profiles and the identification of genes that are important for specific and/or desired strain properties such as pathogenicity, productivity and stress tolerance are facilitated by nucleic acid array methodologies. Also, comparisons of the profile of expression of genes of the invention during the course of a fermentation reaction are possible using nucleic acid array technology.

**Example 13: Analysis of the Dynamics of Cellular Protein Populations (Proteomics)**

The genes, compositions, and methods of the invention may be applied to study the interactions and dynamics of populations of proteins, termed 'proteomics'. Protein populations of interest include, but are not limited to, the total protein population of *C. glutamicum* (e.g., in comparison with the protein populations of other organisms), those proteins which are active under specific environmental or metabolic conditions (e.g., during fermentation, at high or low temperature, or at high or low pH), or those proteins which are active during specific phases of growth and development.

Protein populations can be analyzed by various well-known techniques, such as gel electrophoresis. Cellular proteins may be obtained, for example, by lysis or extraction, and may be separated from one another using a variety of electrophoretic techniques. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins largely on the basis of their molecular weight. Isoelectric focusing polyacrylamide gel electrophoresis (IEF-PAGE) separates proteins by their isoelectric point (which reflects not only the amino acid sequence but also posttranslational modifications of the protein). Another, more preferred method of protein analysis is the consecutive combination of both IEF-PAGE and SDS-PAGE, known as 2-D-gel electrophoresis (described, for example, in Hermann *et al.* (1998) *Electrophoresis* 19: 3217-3221; Fountoulakis *et al.* (1998) *Electrophoresis* 19: 1193-1202; Langen *et al.* (1997) *Electrophoresis* 18: 1184-1192; Antelmann *et al.* (1997) *Electrophoresis* 18: 1451-1463). Other separation techniques may also be utilized for protein separation, such as capillary gel electrophoresis; such techniques are well known in the art.

Proteins separated by these methodologies can be visualized by standard techniques, such as by staining or labeling. Suitable stains are known in the art, and include Coomassie Brilliant Blue, silver stain, or fluorescent dyes such as Sypro Ruby (Molecular Probes). The inclusion of radioactively labeled amino acids or other protein precursors (e.g.,  $^{35}\text{S}$ -methionine,  $^{35}\text{S}$ -cysteine,  $^{14}\text{C}$ -labelled amino acids,  $^{15}\text{N}$ -amino acids) in the medium of *C. glutamicum*

fluorescent labels may be employed. These labeled proteins can be extracted, isolated and separated according to the previously described techniques.

Proteins visualized by these techniques can be further analyzed by measuring the amount of dye or label used. The amount of a given protein can be determined

5 quantitatively using, for example, optical methods and can be compared to the amount of other proteins in the same gel or in other gels. Comparisons of proteins on gels can be made, for example, by optical comparison, by spectroscopy, by image scanning and analysis of gels, or through the use of photographic films and screens. Such techniques are well-known in the art.

10 To determine the identity of any given protein, direct sequencing or other standard techniques may be employed. For example, N- and/or C-terminal amino acid sequencing (such as Edman degradation) may be used, as may mass spectrometry (in particular MALDI or ESI techniques (see, *e.g.*, Langen *et al.* (1997) *Electrophoresis* 18: 1184-1192)). The protein sequences provided herein can be used for the identification  
15 of *C. glutamicum* proteins by these techniques.

The information obtained by these methods can be used to compare patterns of protein presence, activity, or modification between different samples from various biological conditions (*e.g.*, different organisms, time points of fermentation, media conditions, or different biotopes, among others). Data obtained from such experiments  
20 alone, or in combination with other techniques, can be used for various applications, such as to compare the behavior of various organisms in a given (*e.g.*, metabolic) situation, to increase the productivity of strains which produce fine chemicals or to increase the efficiency of the production of fine chemicals.

## 25 Equivalents

Those of ordinary skill in the art will recognize, or will be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the

**TABLE 1: GENES IN THE APPLICATION**

<u>Nucleic Acid</u>	<u>Amino Acid</u>	<u>Identification Code</u>	<u>Contig.</u>	<u>NT Start</u>	<u>NT Stop</u>
<u>SEQ ID NO</u>	<u>SEQ ID NO</u>				
1	2	RXN01638	VV0005	54646	55545
3	4	F RXA01638	GR00456	825	436
5	6	F RXA01639	GR00456	1334	897
7	8	RXN01590	VV0015	66740	60294
9	10	F RXA01590	GR00445	1710	427
11	12	F RXA01542	GR00429	3	5063
13	14	RXN01539	VV0015	60006	57841
15	16	F RXA01539	GR00428	120	2042
17	18	RXN01422	VV0122	17001	16030
19	20	F RXA01422	GR00416	2003	1221
21	22	RXN01403	VV0126	4644	3997
23	24	F RXA01403	GR00409	4410	3997
25	26	RXN01326	VV0102	23923	24288
27	28	F RXA01326	GR00386	45	338
29	30	RXN01301	VV0068	8224	8748
31	32	F RXA01301	GR00375	1993	1589
33	34	RXN01276	VV0009	31190	31633
35	36	F RXA01276	GR00367	29993	30538
37	38	RXN01231	VV0020	30144	29902
39	40	F RXA01231	GR00356	1384	1887
41	42	RXN01210	VV0169	4230	5060
43	44	F RXA01210	GR00349	3	695
45	46	RXN01206	VV0268	303	4
47	48	F RXA01206	GR00346	593	853
49	50	RXN01121	VV0182	5808	6893
51	52	F RXA01121	GR00310	2479	3156
53	54	RXN01085	VV0093	16599	15721
55	56	F RXA01085	GR00303	960	4
57	58	RXN00022	VV0015	27262	28962
59	60	F RXA00022	GR00002	20563	21297
61	62	F RXA01921	GR00551	943	5
63	64	RXN00027	VV0127	60015	59650
65	66	F RXA00027	GR00003	5142	5507
67	68	RXN00028	VV0127	57099	59045
69	70	F RXA00028	GR00003	8058	6112
71	72	RXN00033	VV0127	51753	53087
73	74	RXN00056	VV0044	11980	12729
75	76	F RXA00056	GR00009	1463	714
77	78	RXN00067	VV0019	29740	29255
79	80	F RXA00067	GR00011	708	223
81	82	RXN00077	VV0154	4222	5583
83	84	F RXA00077	GR00012	4228	5589
85	86	RXN00080	VV0154	8446	6917
87	88	F RXA00080	GR00012	7342	6923
89	90	RXN00087	VV0048	204	731
91	92	F RXA00087	GR00013	3983	3456
93	94	RXN00096	VV0129	22302	22000
95	96	F RXA00096	GR00014	4746	5048
97	98	RXN00097	VV0129	21841	20666
99	100	F RXA00097	GR00014	5222	6382

-2-

105	106	RXN00120	VV0142	2612	3451
107	108	F RXA00120	GR00019	2798	3451
109	110	RXN00128	VV0124	7960	9663
111	112	F RXA00128	GR00020	4709	3006
113	114	RXN00154	VV0167	3283	4125
115	116	F RXA00154	GR00023	8568	7726
117	118	RXN00162	VV0084	9489	9842
119	120	F RXA00162	GR00024	5438	5791
121	122	RXN00167	VV0232	4324	4821
123	124	F RXA00167	GR00025	4324	4584
125	126	RXN00171	VV0031	5311	5054
127	128	F RXA00171	GR00026	10316	10086
129	130	RXN00194	VV0115	4174	4614
131	132	F RXA00194	GR00030	290	6
133	134	RXN00197	VV0115	2733	1522
135	136	F RXA00197	GR00030	1731	2741
137	138	RXN00216	VV0096	16292	15303
139	140	RXN00222	VV0096	21079	22224
141	142	F RXA00222	GR00032	21073	22218
143	144	RXN00232	VV0214	601	92
145	146	F RXA00232	GR00035	527	18
147	148	RXN00236	VV0133	3300	2575
149	150	F RXA00236	GR00036	3300	2575
151	152	RXN00242	VV0133	7031	8308
153	154	F RXA00242	GR00036	7031	8233
155	156	RXN00247	VV0057	35082	34156
157	158	F RXA00247	GR00037	7097	6171
159	160	RXN00256	VV0015	3794	4564
161	162	F RXA00256	GR00039	968	1738
163	164	RXN00264	VV0123	14046	13669
165	166	F RXA00264	GR00040	2459	2836
167	168	RXN00267	VV0123	12366	12683
169	170	F RXA00267	GR00040	4091	3822
171	172	RXN00271	VV0019	4975	3986
173	174	F RXA00271	GR00041	3709	2720
175	176	RXN00272	VV0019	5686	6057
177	178	F RXA00272	GR00041	4420	4791
179	180	RXN00283	VV0127	33097	32066
181	182	F RXA00283	GR00045	142	1269
183	184	RXN00334	VV0197	3581	3246
185	186	F RXA00334	GR00057	16762	17097
187	188	RXN00338	VV0197	26797	25658
189	190	F RXA00338	GR00059	1	783
191	192	F RXA00318	GR00055	426	635
193	194	RXN00342	VV0049	1576	1148
195	196	F RXA00342	GR00061	73	501
197	198	RXN00344	VV0135	42719	43597
199	200	F RXA00344	GR00063	6	584
201	202	RXN00353	VV0135	32107	32799
203	204	F RXA00353	GR00068	988	1680
205	206	RXN00354	VV0135	33604	32792



-3-

219	220	RXN00399	VV0025	13735	13421
221	222	F RXA00399	GR00087	830	1144
223	224	RXN00416	VV0181	1	
225	226	F RXA00416	GR00093	1	327
227	228	F RXA00418	GR00094	1	1065
229	230	RXN00422	VV0112	3820	4713
231	232	F RXA00422	GR00097	428	6
233	234	RXN00447	VV0112	23325	22906
235	236	F RXA00447	GR00108	518	817
237	238	RXN00455	VV0076	5523	4774
239	240	F RXA00455	GR00113	2	619
241	242	RXN00473	VV0086	31493	32281
243	244	F RXA00473	GR00119	5799	6563
245	246	RXN00485	VV0086	51200	48906
247	248	F RXA00485	GR00119	25230	23188
249	250	RXN00496	VV0086	17452	18477
251	252	F RXA00496	GR00123	1776	2177
253	254	RXN00503	VV0086	11688	13409
255	256	RXN00504	VV0086	13486	13905
257	258	F RXA00504	GR00125	5007	5252
259	260	RXN00505	VV0086	5810	6304
261	262	F RXA00505	GR00126	1	252
263	264	RXN00507	VV0086	4606	3752
265	266	F RXA00507	GR00127	1098	244
267	268	RXN00510	VV0086	1924	3432
269	270	F RXA00509	GR00128	316	140
271	272	F RXA00510	GR00128	384	914
273	274	RXN00515	VV0144	3289	2588
275	276	F RXA00515	GR00131	3	482
277	278	F RXA00520	GR00132	599	796
279	280	RXN00527	VV0079	23845	25608
281	282	F RXA00527	GR00136	3123	1360
283	284	RXN00547	VV0079	33886	35283
285	286	F RXA00547	GR00142	641	1054
287	288	F RXA00546	GR00142	1	690
289	290	RXN00552	VV0079	27617	28552
291	292	F RXA00552	GR00145	2	718
293	294	RXN00555	VV0079	30437	29499
295	296	F RXA00555	GR00145	2555	1665
297	298	RXN00560	VV0103	7606	7980
299	300	F RXA00560	GR00149	256	492
301	302	RXN00574	VV0323	16133	15255
303	304	F RXA00574	GR00156	767	1645
305	306	RXN00589	VV0323	2680	2231
307	308	F RXA00589	GR00156	14220	14582
309	310	RXN00616	VV0054	5670	5326
311	312	F RXA00616	GR00162	3574	3918
313	314	RXN00647	VV0109	12861	12229
315	316	F RXA00647	GR00169	641	1273
317	318	RXN00653	VV0109	6578	7342
319	320	F RXA00653	GR00169	6924	6160
321	322	RXN00662	VV0142	7711	8979
323	324	F RXA00662	GR00172	2671	1403
325	326	RXN00666	VV0109	1566	2480
327	328	F RXA00666	GR00175	390	4
329	330	RXN00704	VV0005	9240	9866

-4-

333	334	RXN00712	VV0005	1195	500
335	336	F RXA00712	GR00187	1048	500
337	338	RXN00720	VV0232	4899	5564
339	340	F RXA00720	GR00188	7665	7000
341	342	RXN00722	VV0052	2	
343	344	F RXA00722	GR00189	1015	512
345	346	RXN00729	VV0024	3903	2926
347	348	F RXA00729	GR00194	1	642
349	350	F RXA02867	GR10008	610	5
351	352	RXN00730	VV0024	2031	2837
353	354	F RXA00730	GR00194	1063	731
355	356	RXN00731	VV0133	8314	10809
357	358	F RXA00731	GR00195	2613	142
359	360	RXN00738	VV0254	3	
361	362	F RXA00738	GR00201	78	365
363	364	RXN00750	VV0010	30086	30523
365	366	F RXA00750	GR00202	18937	19374
367	368	RXN00762	VV0103	16953	17828
369	370	RXN00768	VV0103	12997	11879
371	372	F RXA00768	GR00204	5956	6399
373	374	F RXA00767	GR00204	5280	5993
375	376	RXN00769	VV0103	11654	11442
377	378	F RXA00769	GR00204	6624	6836
379	380	RXN00771	VV0103	26639	27457
381	382	F RXA00771	GR00205	857	180
383	384	RXN00785	VV0321	2	658
385	386	F RXA00785	GR00207	625	5
387	388	RXN00795	VV0321	6259	5732
389	390	F RXA00795	GR00211	4228	4755
391	392	RXN00831	VV0180	4205	4906
393	394	F RXA00831	GR00224	1662	961
395	396	RXN00835	VV0138	12068	13021
397	398	F RXA00835	GR00226	3	692
399	400	RXN00836	VV0138	13126	14841
401	402	F RXA00836	GR00226	797	2467
403	404	RXN00840	VV0138	6220	6933
405	406	F RXA00840	GR00228	742	1455
407	408	RXN00841	VV0138	6944	7480
409	410	F RXA00841	GR00228	1466	2002
411	412	RXN00846	VV0138	2367	1498
413	414	F RXA00846	GR00230	391	5
415	416	RXN00850	VV0067	4148	3321
417	418	RXN00854	VV0067	371	159
419	420	F RXA00854	GR00231	4708	4920
421	422	RXN00855	VV0255	735	1019
423	424	F RXA00855	GR00232	526	242
425	426	RXN00869	VV0127	5599	4679
427	428	F RXA00869	GR00239	1	792
429	430	RXN00915	VV0238	886	257
431	432	F RXA00915	GR00251	514	5
433	434	RXN00917	VV0238	7202	4524
435	436	F RXA00917	GR00251	5534	4152
437	438	RXN00921	VV0373	1	513
439	440	F RXA00921	GR00252	4750	2852

-5-

447	448	RXN00945	VV0107	1876	2847
449	450	RXN00946	VV0107	3034	3807
451	452	F RXA00946	GR00259	3034	3807
453	454	RXN00953	VV0260	1834	1082
455	456	RXN00959	VV0208	402	857
457	458	F RXA00959	GR00265	402	728
459	460	RXN00963	VV0249	1816	2652
461	462	F RXA00963	GR00269	442	5
463	464	RXN00971	VV0149	14706	14389
465	466	F RXA00971	GR00273	1421	1149
467	468	RXN00991	VV0210	4424	3045
469	470	RXN01004	VV0210	3045	1984
471	472	RXN01016	VV0209	14476	15123
473	474	F RXA01016	GR00290	1141	494
475	476	RXN01023	VV0143	2343	3320
477	478	F RXA01023	GR00292	1817	867
479	480	RXN01028	VV0015	29000	31048
481	482	F RXA01028	GR00295	3	626
483	484	F RXA01812	GR00514	3	1232
485	486	RXN01069	VV0030	986	273
487	488	F RXA01069	GR00299	606	4
489	490	RXN01071	VV0030	4879	2816
491	492	F RXA02898	GR10040	1631	6
493	494	F RXA01071	GR00299	2822	2436
495	496	RXN01075	VV0084	42045	41635
497	498	F RXA01075	GR00300	3269	2859
499	500	RXN01128	VV0157	2427	3440
501	502	F RXA01128	GR00314	1325	312
503	504	RXN01134	VV0077	774	4
505	506	F RXA01134	GR00317	2	460
507	508	RXN01140	VV0077	1642	710
509	510	F RXA01140	GR00318	3272	4057
511	512	RXN01148	VV0136	3147	3746
513	514	F RXA01148	GR00323	1452	2051
515	516	RXN01153	VV0265	546	4
517	518	F RXA01153	GR00325	546	4
519	520	RXN01154	VV0266	644	6
521	522	F RXA01154	GR00326	608	6
523	524	RXN01155	VV0225	252	1721
525	526	F RXA01155	GR00327	1370	6
527	528	RXN01167	VV0117	12777	13172
529	530	F RXA01167	GR00333	3	323
531	532	RXN01169	VV0117	5804	6799
533	534	F RXA01169	GR00334	1	567
535	536	RXN01173	VV0117	11085	10471
537	538	F RXA01173	GR00334	4853	4239
539	540	RXN01174	VV0117	12236	11487
541	542	F RXA01174	GR00334	6004	5255
543	544	RXN01229	VV0020	32482	31205
545	546	F RXA01229	GR00355	2806	3498
547	548	RXN01246	VV0104	2815	3321
549	550	F RXA01246	GR00360	1824	2462
551	552	RXN01249	VV0271	1	
553	554	F RXA01249	GR00363	303	4
555	556	RXN01251	VV0219	13143	12835
		F RXA01251	GR00365	228	536

-6-

561	562	F RXA01263	GR00367	10720	11631
563	564	RXN01266	VV0009	15553	14519
565	566	F RXA01266	GR00367	14457	13423
567	568	RXN01275	VV0009	29514	30431
569	570	F RXA01275	GR00367	28418	29335
571	572	RXN01281	VV0212	4506	5267
573	574	F RXA01281	GR00369	3869	4630
575	576	RXN01296	VV0209	10462	9380
577	578	F RXA01296	GR00373	5836	4754
579	580	RXN01306	VV0148	8158	7151
581	582	F RXA01306	GR00376	5691	4684
583	584	RXN01324	VV0082	6588	6887
585	586	RXN01331	VV0005	39816	42212
587	588	F RXA01331	GR00387	1606	1031
589	590	F RXA00668	GR00176	797	6
591	592	F RXA00674	GR00177	755	6
593	594	RXN01337	VV0032	1925	3337
595	596	F RXA01337	GR00389	5065	3653
597	598	RXN01351	VV0123	2841	3425
599	600	RXN01362	VV0051	27040	23387
601	602	F RXA01362	GR00395	3	1397
603	604	F RXA01364	GR00396	1869	4
605	606	RXN01379	VV0091	10518	11459
607	608	F RXA01379	GR00402	926	6
609	610	RXN01390	VV0277	6654	7310
611	612	F RXA01390	GR00408	992	336
613	614	RXN01391	VV0277	5568	6257
615	616	F RXA01391	GR00408	2078	1389
617	618	RXN01400	VV0126	2988	1489
619	620	F RXA01400	GR00409	2988	1489
621	622	RXN01409	VV0278	5304	4483
623	624	F RXA01409	GR00410	5296	4481
625	626	RXN01434	VV0050	13792	10841
627	628	F RXA01434	GR00417	10228	9863
629	630	RXN01448	VV0089	9602	10768
631	632	F RXA01448	GR00418	19796	19017
633	634	RXN01459	VV0233	3311	4120
635	636	F RXA01459	GR00420	3311	4120
637	638	RXN01460	VV0233	4066	4359
639	640	F RXA01460	GR00420	4066	4359
641	642	RXN01471	VV0019	11467	10661
643	644	F RXA01471	GR00422	5243	4437
645	646	RXN01479	VV0019	18635	18874
647	648	F RXA01479	GR00422	12423	12650
649	650	RXN01484	VV0019	26292	25747
651	652	F RXA01484	GR00422	20068	19523
653	654	RXN01485	VV0019	26454	28505
655	656	F RXA01485	GR00422	20230	22281
657	658	RXN01492	VV0139	36004	36807
659	660	F RXA01492	GR00423	6133	5330
				23238	23711

-7-

675	676	RXN01574	VV0009	48980	47946
677	678	F RXA01574	GR00438	6963	5929
679	680	RXN01589	VV0227	1216	197
681	682	RXN01592	VV0229	14706	13405
683	684	F RXA01592	GR00447	3	1295
685	686	RXN01597	VV0229	8480	7299
687	688	F RXA01597	GR00447	6220	7401
689	690	RXN01598	VV0229	7286	6324
691	692	F RXA01598	GR00447	7414	8376
693	694	RXN01618	VV0050	23629	23246
695	696	F RXA01618	GR00451	1387	1004
697	698	RXN01634	VV0050	43466	42915
699	700	F RXA01634	GR00454	4988	5539
701	702	RXN01635	VV0050	42879	42139
703	704	F RXA01635	GR00454	5575	6315
705	706	RXN01647	VV0005	43276	44445
707	708	F RXA01647	GR00456	12422	11535
709	710	RXN01658	VV0010	44183	42351
711	712	F RXA01658	GR00461	5	1489
713	714	RXN01659	VV0089	5059	5604
715	716	F RXA01659	GR00462	3	488
717	718	RXN01663	VV0089	4271	5128
719	720	F RXA01663	GR00463	438	4
721	722	RXN01669	VV0057	4529	5443
723	724	F RXA01669	GR00465	1002	271
725	726	RXN01672	VV0179	7849	8190
727	728	F RXA01672	GR00467	2	310
729	730	RXN01694	VV0139	13054	13953
731	732	F RXA01694	GR00474	3931	3032
733	734	RXN01696	VV0115	1381	203
735	736	F RXA01696	GR00475	799	203
737	738	RXN01697	VV0139	1581	625
739	740	F RXA01697	GR00476	761	1486
741	742	RXN01701	VV0162	375	4
743	744	F RXA01701	GR00478	196	528
745	746	RXN01703	VV0089	7108	8220
747	748	F RXA01703	GR00479	2118	1648
749	750	RXN01709	VV0022	847	416
751	752	F RXA01709	GR00483	745	416
753	754	RXN01711	VV0191	8153	6996
755	756	F RXA01711	GR00484	2007	850
757	758	RXN01721	VV0036	1026	4
759	760	RXN01734	VV0221	1251	1784
761	762	F RXA01734	GR00492	544	1077
763	764	RXN01742	VV0233	5246	4743
765	766	F RXA01742	GR00493	7614	8117
767	768	RXN01754	VV0127	38790	36850
769	770	F RXA01754	GR00497	4082	2142

-8-

789	790	F RXA01769	GR00501	1275	847
791	792	RXN01771	VV0050	35063	35764
793	794	F RXA01771	GR00502	886	185
795	796	RXN01774	VV0015	1794	2519
797	798	F RXA01774	GR00503	634	1416
799	800	RXN01787	VV0039	256	948
801	802	F RXA01787	GR00506	2	355
803	804	RXN01796	VV0137	2070	2843
805	806	F RXA01796	GR00508	2	484
807	808	RXN01803	VV0216	3355	4314
809	810	F RXA01803	GR00509	5671	4712
811	812	RXN01809	VV0081	9171	10346
813	814	F RXA01062	GR00297	490	5
815	816	F RXA01809	GR00510	3	638
817	818	RXN01811	VV0146	1243	1923
819	820	RXN01813	VV0084	46618	45953
821	822	F RXA01813	GR00515	635	6
823	824	RXN01815	VV0084	49277	50068
825	826	F RXA01815	GR00515	3294	4085
827	828	RXN01825	VV0083	2847	2578
829	830	F RXA01825	GR00516	2847	2578
831	832	RXN01831	VV0083	10874	10413
833	834	F RXA01831	GR00516	10874	10413
835	836	RXN01834	VV0143	11244	11945
837	838	F RXA01834	GR00517	2478	1777
839	840	RXN01846	VV0010	287	6
841	842	F RXA01846	GR00523	261	4
843	844	RXN01847	VV0139	19018	18284
845	846	F RXA01847	GR00524	52	786
847	848	RXN01874	VV0248	352	5
849	850	F RXA01874	GR00535	2556	2903
851	852	RXN01875	VV0145	2894	2049
853	854	F RXA01875	GR00536	516	1313
855	856	F RXA02734	GR00762	6514	6897
857	858	RXN01877	VV0105	3493	2423
859	860	F RXA01877	GR00537	135	1199
861	862	RXN01879	VV0105	1505	573
863	864	F RXA01879	GR00537	2117	2704
865	866	F RXA01880	GR00537	2641	3048
867	868	RXN01896	VV0098	75888	76523
869	870	F RXA01896	GR00544	2	580
871	872	RXN01899	VV0098	77817	78602
873	874	F RXA01899	GR00544	1874	2659
875	876	RXN01902	VV0098	84095	83037
877	878	F RXA01902	GR00544	7957	7094
879	880	RXN01908	VV0187	4030	4875
881	882	F RXA01908	GR00545	4030	4512
883	884	RXN01909	VV0218	69	947
885	886	F RXA01909	GR00546	59	937
887	888	RXN01910	VV0218	1040	1885
889	890	F RXA01910	GR00546	1030	1875
891	892	RXN01911	VV0218	2209	3147
893	894	F RXA01911	GR00546	2199	3044
895	896	RXN01930	VV0127	46545	47495
897	898	F RXA01930	GR00555	3817	2867
899	900	RXN01944	VV0050	42128	41157

-9-

903	904	F RXA01636	GR00454	6326	6898
905	906	RXN01945	VV0050	41150	39159
907	908	F RXA01945	GR00558	392	1633
909	910	F RXA01627	GR00453	1	495
911	912	RXN01960	VV0200	2259	1942
913	914	F RXA01960	GR00565	187	504
915	916	RXN01985	VV0056	1331	282
917	918	RXN01987	VV0149	167	379
919	920	F RXA01987	GR00576	167	379
921	922	RXN01988	VV0149	887	462
923	924	F RXA01988	GR00576	779	462
925	926	RXN01991	VV0230	926	1798
927	928	F RXA01991	GR00581	926	1720
929	930	RXN01996	VV0174	28434	27898
931	932	F RXA01996	GR00585	88	624
933	934	RXN02007	VV0324	855	223
935	936	F RXA02007	GR00598	651	223
937	938	RXN02014	VV0137	8298	8804
939	940	F RXA02014	GR00607	935	540
941	942	RXN02019	VV0129	44705	44205
943	944	F RXA02019	GR00612	597	106
945	946	RXN02023	VV0160	3234	4001
947	948	F RXA02023	GR00613	3234	4001
949	950	RXN02032	VV0117	5181	5750
951	952	F RXA02032	GR00618	4160	4729
953	954	RXN02039	VV0190	1482	643
955	956	F RXA02039	GR00621	3	812
957	958	RXN02044	VV0025	17208	15826
959	960	RXN02045	VV0025	15823	15563
961	962	F RXA02045	GR00623	1913	2173
963	964	RXN02049	VV0009	35549	36157
965	966	F RXA02049	GR00624	1583	2029
967	968	RXN02050	VV0009	36003	36797
969	970	F RXA02050	GR00624	2462	2833
971	972	RXN02059	VV0222	10306	10800
973	974	F RXA02059	GR00625	4678	4184
975	976	RXN02066	VV0222	6187	6678
977	978	F RXA02066	GR00626	6187	6678
979	980	RXN02067	VV0222	6733	7188
981	982	F RXA02067	GR00626	6733	7188
983	984	RXN02075	VV0318	12990	13778
985	986	RXN02076	VV0318	13879	14412
987	988	F RXA02076	GR00628	6902	7435
989	990	RXN02094	VV0126	18268	18984
991	992	F RXA02094	GR00629	13282	13998
993	994	RXN02104	VV0318	7435	6314
995	996	F RXA02104	GR00631	5327	4908
997	998	F RXA02071	GR00628	458	6
999	1000	RXN02107	VV0123	21585	21244
1001	1002	F RXA02107	GR00632	1536	1877
1003	1004	RXN02108	VV0123	21217	20609
1005	1006	F RXA02108	GR00632	2077	2511
1007	1008	RXN02114	VV0180	3	
1009	1010	F RXA02114	GR00634	615	130
1011	1012	RXN02121	VV0102	12833	12129
1013	1014	F RXA02121	GR00636	5813	5109
				4109	4750

-10-

1017	1018	F RXA02138	GR00639	4409	4750
1019	1020	RXN02151	VV0300	19913	21100
1021	1022	F RXA02151	GR00639	19913	21100
1023	1024	RXN02169	VV0100	3172	4017
1025	1026	F RXA02169	GR00641	3172	4017
1027	1028	RXN02180	VV0100	16813	15356
1029	1030	F RXA02180	GR00641	16813	15356
1031	1032	RXN02185	VV0100	20185	20763
1033	1034	F RXA02185	GR00641	20185	20763
1035	1036	RXN02186	VV0100	21192	20995
1037	1038	F RXA02186	GR00641	21213	20995
1039	1040	RXN02207	VV0302	802	5
1041	1042	F RXA02207	GR00646	10909	11667
1043	1044	RXN02223	VV0308	2732	3232
1045	1046	F RXA02223	GR00652	425	6
1047	1048	RXN02226	VV0068	1059	4
1049	1050	F RXA02226	GR00653	1059	4
1051	1052	RXN02238	VV0204	1345	1629
1053	1054	F RXA02238	GR00654	5241	5525
1055	1056	RXN02254	VV0202	2	
1057	1058	F RXA02254	GR00654	21769	22449
1059	1060	RXN02271	VV0020	14281	14838
1061	1062	F RXA02271	GR00655	5406	5963
1063	1064	RXN02279	VV0020	236	1693
1065	1066	F RXA02279	GR00657	1	1404
1067	1068	RXN02296	VV0127	24138	24626
1069	1070	F RXA02296	GR00662	6978	7466
1071	1072	RXN02300	VV0127	28354	28022
1073	1074	F RXA02300	GR00662	11194	10862
1075	1076	RXN02301	VV0127	29070	28354
1077	1078	F RXA02301	GR00662	11910	11194
1079	1080	RXN02302	VV0127	29196	30074
1081	1082	F RXA02302	GR00662	12036	12800
1083	1084	RXN02303	VV0127	13326	14231
1085	1086	F RXA02303	GR00663	1	720
1087	1088	RXN02307	VV0127	12611	11991
1089	1090	F RXA02307	GR00664	395	6
1091	1092	RXN02314	VV0025	23092	23532
1093	1094	F RXA02314	GR00665	6379	5939
1095	1096	RXN02337	VV0141	4679	3357
1097	1098	F RXA02337	GR00672	2893	3816
1099	1100	RXN02339	VV0195	1	
1101	1102	F RXA02339	GR00674	1	492
1103	1104	RXN02340	VV0195	1640	576
1105	1106	F RXA02338	GR00673	484	5
1107	1108	F RXA02340	GR00674	1214	576
1109	1110	RXN02341	VV0078	4279	4764
1111	1112	F RXA02341	GR00675	415	5
1113	1114	RXN02360	VV0051	14638	12206
1115	1116	F RXA02360	GR00685	3644	6076
1117	1118	RXN02361	VV0051	12122	11472
1119	1120	F RXA02361	GR00685	6160	6810
1121	1122	RXN02367	VV0102	4639	5247
1123	1124	F RXA02367	GR00687	2162	1554
1125	1126	RXN02368	VV0102	3883	4557
1127	1128	F RXA02368	GR00687	2918	2244



-11-

1131	1132	F RXA02381	GR00691	1792	770
1133	1134	RXN02383	VV0213	639	4
1135	1136	F RXA02383	GR00692	608	6
1137	1138	RXN02387	VV0176	2729	3490
1139	1140	F RXA02387	GR00694	683	6
1141	1142	RXN02398	VV0176	12750	11149
1143	1144	F RXA02398	GR00698	2841	4370
1145	1146	RXN02406	VV0084	22016	22564
1147	1148	F RXA02406	GR00701	1322	774
1149	1150	RXN02407	VV0084	21758	21387
1151	1152	F RXA02407	GR00701	1580	1885
1153	1154	RXN02408	VV0084	20832	19921
1155	1156	F RXA02408	GR00702	832	5
1157	1158	RXN02409	VV0084	21371	20835
1159	1160	F RXA02409	GR00702	1248	835
1161	1162	RXN02428	VV0110	4585	3452
1163	1164	F RXA02428	GR00707	4585	3452
1165	1166	RXN02454	VV0196	2810	1569
1167	1168	F RXA02454	GR00711	3	815
1169	1170	RXN02457	VV0124	19193	18084
1171	1172	F RXA02457	GR00712	1295	2404
1173	1174	RXN02460	VV0124	14649	15152
1175	1176	F RXA02460	GR00712	5839	5336
1177	1178	RXN02464	VV0211	1990	3189
1179	1180	F RXA02464	GR00713	1107	1613
1181	1182	RXN02465	VV0211	3590	3192
1183	1184	F RXA02465	GR00713	2014	1616
1185	1186	RXN02466	VV0211	92	6
1187	1188	F RXA02466	GR00714	92	6
1189	1190	RXN02505	VV0007	23969	24139
1191	1192	F RXA02505	GR00720	18423	18593
1193	1194	RXN02510	VV0171	17467	16832
1195	1196	F RXA02510	GR00721	1983	2618
1197	1198	RXN02519	VV0183	2709	4337
1199	1200	F RXA02519	GR00724	1933	128
1201	1202	RXN02520	VV0183	2243	1560
1203	1204	F RXA02520	GR00724	2222	2905
1205	1206	RXN02534	VV0057	11192	11995
1207	1208	F RXA02534	GR00726	5536	6339
1209	1210	RXN02537	VV0057	14617	15078
1211	1212	F RXA02537	GR00726	8961	9422
1213	1214	RXN02538	VV0057	15078	15749
1215	1216	F RXA02538	GR00726	9422	10093
1217	1218	RXN02555	VV0101	5340	4738
1219	1220	F RXA02555	GR00731	1757	1155
1221	1222	RXN02564	VV0154	10016	9015
1223	1224	F RXA02564	GR00732	2543	3217
1225	1226	RXN02568	VV0245	1657	5
1227	1228	F RXA02568	GR00735	1363	5
1229	1230	RXN02593	VV0098	11073	11669
1231	1232	F RXA02593	GR00741	18693	18481
1233	1234	F RXA02594	GR00741	19077	18754
1235	1236	RXN02606	VV0098	34557	35927
1237	1238	F RXA02606	GR00742	13514	12144
1239	1240	RXN02610	VV0098	31620	30694
		F RXA02610	GR00742	16452	17378

1131 1132 F RXA02381 GR00691 1792 770  
 1133 1134 RXN02383 VV0213 639 4  
 1135 1136 F RXA02383 GR00692 608 6  
 1137 1138 RXN02387 VV0176 2729 3490  
 1139 1140 F RXA02387 GR00694 683 6  
 1141 1142 RXN02398 VV0176 12750 11149  
 1143 1144 F RXA02398 GR00698 2841 4370  
 1145 1146 RXN02406 VV0084 22016 22564  
 1147 1148 F RXA02406 GR00701 1322 774  
 1149 1150 RXN02407 VV0084 21758 21387  
 1151 1152 F RXA02407 GR00701 1580 1885  
 1153 1154 RXN02408 VV0084 20832 19921  
 1155 1156 F RXA02408 GR00702 832 5  
 1157 1158 RXN02409 VV0084 21371 20835  
 1159 1160 F RXA02409 GR00702 1248 835  
 1161 1162 RXN02428 VV0110 4585 3452  
 1163 1164 F RXA02428 GR00707 4585 3452  
 1165 1166 RXN02454 VV0196 2810 1569  
 1167 1168 F RXA02454 GR00711 3 815  
 1169 1170 RXN02457 VV0124 19193 18084  
 1171 1172 F RXA02457 GR00712 1295 2404  
 1173 1174 RXN02460 VV0124 14649 15152  
 1175 1176 F RXA02460 GR00712 5839 5336  
 1177 1178 RXN02464 VV0211 1990 3189  
 1179 1180 F RXA02464 GR00713 1107 1613  
 1181 1182 RXN02465 VV0211 3590 3192  
 1183 1184 F RXA02465 GR00713 2014 1616  
 1185 1186 RXN02466 VV0211 92 6  
 1187 1188 F RXA02466 GR00714 92 6  
 1189 1190 RXN02505 VV0007 23969 24139  
 1191 1192 F RXA02505 GR00720 18423 18593  
 1193 1194 RXN02510 VV0171 17467 16832  
 1195 1196 F RXA02510 GR00721 1983 2618  
 1197 1198 RXN02519 VV0183 2709 4337  
 1199 1200 F RXA02519 GR00724 1933 128  
 1201 1202 RXN02520 VV0183 2243 1560  
 1203 1204 F RXA02520 GR00724 2222 2905  
 1205 1206 RXN02534 VV0057 11192 11995  
 1207 1208 F RXA02534 GR00726 5536 6339  
 1209 1210 RXN02537 VV0057 14617 15078  
 1211 1212 F RXA02537 GR00726 8961 9422  
 1213 1214 RXN02538 VV0057 15078 15749  
 1215 1216 F RXA02538 GR00726 9422 10093  
 1217 1218 RXN02555 VV0101 5340 4738  
 1219 1220 F RXA02555 GR00731 1757 1155  
 1221 1222 RXN02564 VV0154 10016 9015  
 1223 1224 F RXA02564 GR00732 2543 3217  
 1225 1226 RXN02568 VV0245 1657 5  
 1227 1228 F RXA02568 GR00735 1363 5  
 1229 1230 RXN02593 VV0098 11073 11669  
 1231 1232 F RXA02593 GR00741 18693 18481  
 1233 1234 F RXA02594 GR00741 19077 18754  
 1235 1236 RXN02606 VV0098 34557 35927  
 1237 1238 F RXA02606 GR00742 13514 12144  
 1239 1240 RXN02610 VV0098 31620 30694  
 F RXA02610 GR00742 16452 17378

-12-

1245	1246	F RXA02624	GR00746	5602	4889
1247	1248	RXN02626	VV0314	2012	1008
1249	1250	RXN02656	VV0090	15756	14917
1251	1252	RXN02673	VV0315	14030	13398
1253	1254	F RXA02673	GR00753	14030	13398
1255	1256	RXN02680	VV0098	64917	66200
1257	1258	F RXA02680	GR00754	6392	5109
1259	1260	F RXA02679	GR00754	5268	5693
1261	1262	F RXA02681	GR00754	5751	6194
1263	1264	RXN02693	VV0098	74100	75875
1265	1266	F RXA02693	GR00755	1650	4
1267	1268	RXN02696	VV0017	7946	7491
1269	1270	F RXA02696	GR00756	742	287
1271	1272	RXN02697	VV0017	31257	32783
1273	1274	F RXA02697	GR00757	1	699
1275	1276	F RXA02719	GR00758	19598	20245
1277	1278	RXN02720	VV0017	8727	8026
1279	1280	F RXA02720	GR00759	631	5
1281	1282	RXN02744	VV0074	893	1696
1283	1284	F RXA02744	GR00763	14460	13657
1285	1286	RXN02770	VV0171	4202	2637
1287	1288	F RXA02770	GR00772	3	1322
1289	1290	RXN02781	VV0084	14566	13376
1291	1292	F RXA02781	GR00774	1345	155
1293	1294	RXN02782	VV0093	7148	8446
1295	1296	F RXA02782	GR00775	204	875
1297	1298	RXN02812	VV0210	342	4
1299	1300	F RXA02812	GR00793	2	568
1301	1302	RXN02817	VV0346	403	5
1303	1304	F RXA02817	GR00798	403	5
1305	1306	RXN02818	VV0347	611	6
1307	1308	F RXA02818	GR00799	611	6
1309	1310	RXN02825	VV0082	3589	1751
1311	1312	F RXA01322	GR00385	443	6
1313	1314	F RXA02824	GR00805	531	4
1315	1316	F RXA02825	GR00806	565	182
1317	1318	RXN02838	VV0161	1	
1319	1320	F RXA02838	GR00831	1	462
1321	1322	RXN02840	VV0365	488	339
1323	1324	F RXA02840	GR00835	488	339
1325	1326	RXN02841	VV0055	11788	12222
1327	1328	F RXA02841	GR00840	283	5
1329	1330	RXN02846	VV0127	30861	30112
1331	1332	F RXA02846	GR00845	578	6
1333	1334	RXN02847	VV0113	47	1135
1335	1336	F RXA02847	GR00847	598	5
1337	1338	RXN02849	VV0237	2	
1339	1340	F RXA02849	GR00849	2	283
1341	1342	RXN02911	VV0135	24643	25101
1343	1344	RXN02914	VV0127	17305	16763
1345	1346	RXN02921	VV0213	1871	1401
1347	1348	RXN02924	VV0088	4557	5105
1349	1350	RXN02927	VV0082	18836	19303
1351	1352	RXN02928	VV0082	19511	20203
1353	1354	RXN02931	VV0090	25420	25644
1355	1356	RXN02932	VV0176	23391	24362

1359	1360	RXN02936	VV0197	24360	24557
1361	1362	RXN02939	VV0008	33988	32387
1363	1364	F RXA01383	GR00406	1147	5
1365	1366	RXN02950	VV0224	7629	7306
1367	1368	RXN02951	VV0176	5739	5131
1369	1370	RXN02957	VV0020	30448	30158
1371	1372	RXN02967	VV0318	6614	6931
1373	1374	RXN02971	VV0210	951	640
1375	1376	RXN02978	VV0010	2191	683
1377	1378	RXN02995	VV0069	348	1913
1379	1380	RXN02997	VV0069	3709	2981
1381	1382	RXN03001	VV0170	422	874
1383	1384	RXN03005	VV0237	1101	334
1385	1386	RXN03009	VV0238	353	6
1387	1388	RXN03010	VV0238	7435	7199
1389	1390	RXN03011	VV0098	2984	3184
1391	1392	RXN03012	VV0241	2	571
1393	1394	RXN03017	VV0218	5720	7258
1395	1396	F RXA02753	GR00765	2630	138
1397	1398	RXN03018	VV0218	7221	8213
1399	1400	RXN03024	VV0003	6315	7730
1401	1402	RXN03025	VV0003	8668	7796
1403	1404	RXN03027	VV0008	17	151
1405	1406	RXN03029	VV0009	95	607
1407	1408	RXN03031	VV0011	1	789
1409	1410	RXN03032	VV0012	3652	3936
1411	1412	RXN03034	VV0013	2	661
1413	1414	F RXA00063	GR00010	1658	1374
1415	1416	RXN03037	VV0015	5364	5549
1417	1418	RXN03041	VV0018	1770	1273
1419	1420	F RXA02892	GR10035	1171	668
1421	1422	RXN03045	VV0019	33044	34039
1423	1424	RXN03046	VV0020	1	336
1425	1426	RXN03047	VV0020	25070	26485
1427	1428	F RXA00036	GR00004	7204	8619
1429	1430	RXN03048	VV0020	27423	26551
1431	1432	F RXA00037	GR00004	9557	8685
1433	1434	RXN03050	VV0021	6368	7333
1435	1436	RXN03053	VV0026	12	1535
1437	1438	F RXA02885	GR10021	1	1536
1439	1440	RXN03055	VV0026	3519	3947
1441	1442	F RXA00261	GR00039	11693	11265
1443	1444	F RXA02888	GR10024	326	754
1445	1446	RXN03059	VV0030	5373	4894
1447	1448	F RXA02899	GR10040	2125	1646
1449	1450	RXN03062	VV0035	525	4
1451	1452	RXN03066	VV0038	7298	6636
1453	1454	F RXA02876	GR10016	405	1067
1455	1456	RXN03067	VV0038	7493	7323
1457	1458	RXN03068	VV0038	7648	7529
1459	1460	RXN03073	VV0042	1573	944
1461	1462	F RXA02905	GR10044	477	4
1463	1464	RXN03085	VV0048	4511	4161
1465	1466	RXN03089	VV0053	1183	26
1467	1468	F RXA00071	GR00011	4013	5464
		RXN03098	VV0064	2100	2723

-14-

1473	1474	RXN03104	VV0071	401	577
1475	1476	RXN03106	VV0074	15930	16121
1477	1478	RXN03107	VV0076	232	432
1479	1480	RXN03113	VV0086	6541	8139
1481	1482	F RXA00506	GR00126	489	1829
1483	1484	RXN03115	VV0089	148	546
1485	1486	RXN03122	VV0104	3329	3475
1487	1488	RXN03134	VV0127	65312	65662
1489	1490	RXN03135	VV0127	66674	67402
1491	1492	F RXA02285	GR00660	1544	2272
1493	1494	RXN03138	VV0129	21194	21664
1495	1496	RXN03140	VV0131	4550	4302
1497	1498	RXN03141	VV0135	31144	31473
1499	1500	RXN03146	VV0143	25998	26468
1501	1502	RXN03147	VV0144	2726	2977
1503	1504	RXN03149	VV0146	969	1235
1505	1506	RXN03152	VV0166	264	536
1507	1508	RXN03153	VV0176	46481	47044
1509	1510	RXN03154	VV0179	1328	2239
1511	1512	RXN03156	VV0187	4908	5087
1513	1514	F RXA00176	GR00027	3475	3317
1515	1516	RXN03162	VV0195	909	1304
1517	1518	RXN03167	VV0327	633	4
1519	1520	F RXA02862	GR10006	1695	2330
1521	1522	RXN03170	VV0328	457	209
1523	1524	F RXA02856	GR10003	459	211
1525	1526	RXN03172	VV0329	1392	367
1527	1528	F RXA02858	GR10004	1392	367
1529	1530	RXN03173	VV0330	1340	243
1531	1532	F RXA02874	GR10015	1348	869
1533	1534	RXN03174	VV0331	461	6
1535	1536	F RXA02884	GR10020	1695	2156
1537	1538	RXN03177	VV0333	816	151
1539	1540	F RXA02881	GR10019	94	759
1541	1542	RXN03182	VV0339	276	4
1543	1544	RXN03184	VV0374	517	20
1545	1546	RXN03185	VV0375	311	123
1547	1548	RXA00003	GR00001	2279	3019
1549	1550	RXA00008	GR00002	606	115
1551	1552	RXA00015	GR00002	5999	6307
1553	1554	RXA00018	GR00002	12979	14277
1555	1556	RXA00020	GR00002	17142	16363
1557	1558	RXA00021	GR00002	18766	20538
1559	1560	RXA00025	GR00003	2211	3647
1561	1562	RXA00031	GR00003	10383	9982
1563	1564	RXA00049	GR00008	2270	2956
1565	1566	RXA00052	GR00008	7957	7247
1567	1568	RXA00054	GR00008	8557	11469
1569	1570	RXA00058	GR00009	7394	6831
1571	1572	RXA00059	GR00009	8301	8020
1573	1574	RXA00065	GR00010	4140	4412
1575	1576	RXA00068	GR00011	1305	724
1577	1578	RXA00079	GR00012	6599	6820
1579	1580	RXA00082	GR00012	9019	8456
1581	1582	RXA00083	GR00013	771	1070
			GR00014	204	2426

-15-

1587	1588	RXA00108	GR00015	546	4
1589	1590	RXA00110	GR00016	364	912
1591	1592	RXA00117	GR00019	791	201
1593	1594	RXA00118	GR00019	918	1172
1595	1596	RXA00119	GR00019	1704	2462
1597	1598	RXA00121	GR00019	3473	4183
1599	1600	RXA00122	GR00019	4220	5842
1601	1602	RXA00127	GR00020	2871	2416
1603	1604	RXA00134	GR00021	1648	1079
1605	1606	RXA00140	GR00022	3841	3656
1607	1608	RXA00141	GR00022	4307	3846
1609	1610	RXA00142	GR00022	4776	4300
1611	1612	RXA00150	GR00023	4085	4858
1613	1614	RXA00151	GR00023	4956	5552
1615	1616	RXA00153	GR00023	7656	7231
1617	1618	RXA00155	GR00023	8615	9397
1619	1620	RXA00159	GR00024	3868	2687
1621	1622	RXA00161	GR00024	4893	5354
1623	1624	RXA00169	GR00026	5222	3150
1625	1626	RXA00170	GR00026	9914	8061
1627	1628	RXA00173	GR00027	1716	1384
1629	1630	RXA00174	GR00027	2079	1795
1631	1632	RXA00175	GR00027	2732	2103
1633	1634	RXA00179	GR00028	1714	1256
1635	1636	RXA00180	GR00028	2334	1795
1637	1638	RXA00183	GR00028	7344	8195
1639	1640	RXA00185	GR00028	9418	12045
1641	1642	RXA00199	GR00031	2172	754
1643	1644	RXA00200	GR00031	2837	2535
1645	1646	RXA00207	GR00032	6430	6747
1647	1648	RXA00211	GR00032	10120	10782
1649	1650	RXA00218	GR00032	18104	19243
1651	1652	RXA00220	GR00032	20666	20163
1653	1654	RXA00230	GR00034	746	27
1655	1656	RXA00233	GR00036	420	4
1657	1658	RXA00234	GR00036	998	459
1659	1660	RXA00237	GR00036	3668	4045
1661	1662	RXA00238	GR00036	4186	4554
1663	1664	RXA00239	GR00036	5118	4534
1665	1666	RXA00240	GR00036	5342	5133
1667	1668	RXA00244	GR00037	1565	930
1669	1670	RXA00245	GR00037	3049	1565
1671	1672	RXA00248	GR00037	7843	7121
1673	1674	RXA00250	GR00038	6	221
1675	1676	RXA00252	GR00038	485	727
1677	1678	RXA00257	GR00039	1760	2215
1679	1680	RXA00258	GR00039	3219	3890
1681	1682	RXA00260	GR00039	9234	10409
1683	1684	RXA00273	GR00042	185	1297
1685	1686	RXA00274	GR00042	1556	4165
1687	1688	RXA00275	GR00042	4696	4238
1689	1690	RXA00276	GR00042	5016	4675
1691	1692	RXA00279	GR00043	4001	2616

-16-

1701	1702	RXA00297	GR00048	2861	3772
1703	1704	RXA00320	GR00057	358	537
1705	1706	RXA00321	GR00057	2411	597
1707	1708	RXA00322	GR00057	3658	2555
1709	1710	RXA00325	GR00057	8594	9238
1711	1712	RXA00326	GR00057	9378	9857
1713	1714	RXA00336	GR00057	19461	19931
1715	1716	RXA00337	GR00058	530	6
1717	1718	RXA00339	GR00059	817	1533
1719	1720	RXA00349	GR00066	3	1061
1721	1722	RXA00355	GR00069	635	510
1723	1724	RXA00375	GR00080	549	49
1725	1726	RXA00380	GR00082	836	216
1727	1728	RXA00387	GR00084	1403	591
1729	1730	RXA00392	GR00086	3890	3027
1731	1732	RXA00394	GR00086	5322	4990
1733	1734	RXA00395	GR00086	5417	5716
1735	1736	RXA00396	GR00086	6653	6183
1737	1738	RXA00397	GR00086	7206	6667
1739	1740	RXA00398	GR00087	1	681
1741	1742	RXA00408	GR00091	642	1088
1743	1744	RXA00409	GR00091	1088	2500
1745	1746	RXA00411	GR00092	1685	1011
1747	1748	RXA00423	GR00097	909	457
1749	1750	RXA00424	GR00097	1379	909
1751	1752	RXA00425	GR00097	1433	1657
1753	1754	RXA00428	GR00098	2657	2025
1755	1756	RXA00429	GR00098	3063	2662
1757	1758	RXA00430	GR00098	3473	3063
1759	1760	RXA00433	GR00100	1446	1970
1761	1762	RXA00451	GR00110	816	325
1763	1764	RXA00457	GR00114	1451	372
1765	1766	RXA00462	GR00116	3023	1644
1767	1768	RXA00463	GR00116	4209	3388
1769	1770	RXA00468	GR00118	1282	464
1771	1772	RXA00469	GR00119	1647	472
1773	1774	RXA00472	GR00119	5449	4589
1775	1776	RXA00474	GR00119	6575	8152
1777	1778	RXA00475	GR00119	8822	8163
1779	1780	RXA00476	GR00119	8961	9821
1781	1782	RXA00481	GR00119	17636	18220
1783	1784	RXA00486	GR00120	1	702
1785	1786	RXA00490	GR00121	2676	1774
1787	1788	RXA00491	GR00122	1057	638
1789	1790	RXA00493	GR00123	3	326
1791	1792	RXA00519	GR00132	4	516
1793	1794	RXA00528	GR00136	3562	4650
1795	1796	RXA00529	GR00136	5274	4732
1797	1798	RXA00530	GR00136	6837	5557
1799	1800	RXA00535	GR00137	5155	5871
1801	1802	RXA00540	GR00139	2027	2269
1803	1804	RXA00549	GR00143	502	897
1805	1806	RXA00550	GR00143	935	1255
1807	1808	RXA00553	GR00145	742	1062
1809	1810	RXA00554	GR00145	1606	1136
1811	1812	RXA00563	GR00151	1	2739
1813	1814	RXA00564	GR00151	3744	4148

-17-

1815	1816	RXA00573	GR00156	117	767
1817	1818	RXA00576	GR00156	2916	2245
1819	1820	RXA00577	GR00156	2980	3327
1821	1822	RXA00578	GR00156	4087	3365
1823	1824	RXA00582	GR00156	9442	8924
1825	1826	RXA00585	GR00156	11894	11577
1827	1828	RXA00586	GR00156	12818	11937
1829	1830	RXA00587	GR00156	13008	13490
1831	1832	RXA00595	GR00159	3	332
1833	1834	RXA00597	GR00159	797	1066
1835	1836	RXA00598	GR00159	1070	1387
1837	1838	RXA00601	GR00159	3459	3749
1839	1840	RXA00602	GR00159	4907	4155
1841	1842	RXA00604	GR00159	5489	5779
1843	1844	RXA00610	GR00161	1193	2056
1845	1846	RXA00611	GR00161	3640	2165
1847	1848	RXA00613	GR00162	1652	1200
1849	1850	RXA00614	GR00162	1680	2594
1851	1852	RXA00617	GR00162	4002	5084
1853	1854	RXA00628	GR00165	1284	877
1855	1856	RXA00631	GR00166	172	1626
1857	1858	RXA00637	GR00167	2002	2754
1859	1860	RXA00646	GR00169	446	6
1861	1862	RXA00649	GR00169	2823	3278
1863	1864	RXA00652	GR00169	5449	5997
1865	1866	RXA00654	GR00169	7213	8478
1867	1868	RXA00656	GR00169	9495	9235
1869	1870	RXA00657	GR00169	10882	9980
1871	1872	RXA00661	GR00172	664	1353
1873	1874	RXA00667	GR00175	593	1177
1875	1876	RXA00676	GR00178	647	1393
1877	1878	RXA00678	GR00179	1037	303
1879	1880	RXA00691	GR00181	2152	1223
1881	1882	RXA00692	GR00181	3450	2317
1883	1884	RXA00693	GR00181	4303	3821
1885	1886	RXA00701	GR00182	427	801
1887	1888	RXA00707	GR00185	377	1348
1889	1890	RXA00713	GR00188	71	1033
1891	1892	RXA00714	GR00188	1809	1249
1893	1894	RXA00716	GR00188	3002	3514
1895	1896	RXA00719	GR00188	5283	6911
1897	1898	RXA00724	GR00191	811	164
1899	1900	RXA00726	GR00192	841	701
1901	1902	RXA00740	GR00202	1646	1068
1903	1904	RXA00741	GR00202	2986	2054
1905	1906	RXA00742	GR00202	5517	3868
1907	1908	RXA00743	GR00202	6652	6230
1909	1910	RXA00745	GR00202	13874	13341
1911	1912	RXA00746	GR00202	13755	14945
1913	1914	RXA00747	GR00202	15067	15654
1915	1916	RXA00748	GR00202	15917	16360
1917	1918	RXA00749	GR00202	17240	16542
1919	1920	RXA00751	GR00202	20245	19418
1921	1922	RXA00752	GR00202	21847	21419
1923	1924	RXA00757	GR00203	3119	4372
1925	1926	RXA00763	GR00204	1384	2166
			GR00204	3283	3969

-18-

1929	1930	RXA00781	GR00206	2682	2395
1931	1932	RXA00788	GR00209	910	686
1933	1934	RXA00804	GR00215	438	881
1935	1936	RXA00805	GR00215	2057	2938
1937	1938	RXA00808	GR00217	1029	352
1939	1940	RXA00812	GR00219	287	1345
1941	1942	RXA00814	GR00219	2463	3236
1943	1944	RXA00815	GR00219	3236	3808
1945	1946	RXA00816	GR00219	4382	4678
1947	1948	RXA00826	GR00223	567	37
1949	1950	RXA00830	GR00224	266	988
1951	1952	RXA00853	GR00231	3775	3173
1953	1954	RXA00861	GR00235	6	431
1955	1956	RXA00862	GR00236	580	17
1957	1958	RXA00874	GR00241	758	1846
1959	1960	RXA00876	GR00241	4208	2454
1961	1962	RXA00881	GR00242	8057	8434
1963	1964	RXA00882	GR00242	8788	9465
1965	1966	RXA00883	GR00242	10060	9542
1967	1968	RXA00887	GR00242	13544	14266
1969	1970	RXA00889	GR00242	15341	15928
1971	1972	RXA00893	GR00244	789	193
1973	1974	RXA00895	GR00244	2578	1988
1975	1976	RXA00904	GR00246	1457	702
1977	1978	RXA00908	GR00247	1611	2168
1979	1980	RXA00916	GR00251	4108	518
1981	1982	RXA00926	GR00253	466	104
1983	1984	RXA00930	GR00253	3841	3089
1985	1986	RXA00932	GR00253	5068	5541
1987	1988	RXA00933	GR00253	6047	5586
1989	1990	RXA00940	GR00257	129	524
1991	1992	RXA00949	GR00259	5400	6047
1993	1994	RXA00969	GR00273	1	147
1995	1996	RXA00973	GR00274	2272	1670
1997	1998	RXA00978	GR00276	217	831
1999	2000	RXA00986	GR00280	60	401
2001	2002	RXA00987	GR00280	875	411
2003	2004	RXA00988	GR00280	1371	949
2005	2006	RXA01005	GR00286	520	1365
2007	2008	RXA01007	GR00287	2572	866
2009	2010	RXA01008	GR00287	2719	4659
2011	2012	RXA01011	GR00288	2089	857
2013	2014	RXA01017	GR00290	2175	1567
2015	2016	RXA01021	GR00291	1759	2280
2017	2018	RXA01029	GR00295	1338	1826
2019	2020	RXA01031	GR00295	3182	3847
2021	2022	RXA01032	GR00295	3974	4348
2023	2024	RXA01033	GR00295	4363	4698
2025	2026	RXA01034	GR00295	5177	4824
2027	2028	RXA01035	GR00295	5818	6423
2029	2030	RXA01036	GR00295	6513	6965
2031	2032	RXA01037	GR00295	7000	7527
2033	2034	RXA01038	GR00295	7530	8276
2035	2036	RXA01039	GR00295	9540	8965
2037	2038	RXA01040	GR00295	9711	10613
2039	2040	RXA01041	GR00295	10780	10932
				11088	12365



-19-

2043	2044	RXA01043	GR00295	12774	13346
2045	2046	RXA01044	GR00295	14024	15280
2047	2048	RXA01045	GR00295	15407	17230
2049	2050	RXA01046	GR00295	17441	19219
2051	2052	RXA01047	GR00295	19244	19717
2053	2054	RXA01058	GR00296	8566	8246
2055	2056	RXA01063	GR00297	828	499
2057	2058	RXA01066	GR00298	605	1330
2059	2060	RXA01068	GR00298	2184	3254
2061	2062	RXA01074	GR00300	2811	2107
2063	2064	RXA01076	GR00300	4374	3355
2065	2066	RXA01078	GR00300	6043	6876
2067	2068	RXA01083	GR00302	1777	1502
2069	2070	RXA01088	GR00304	3083	1902
2071	2072	RXA01091	GR00305	546	76
2073	2074	RXA01092	GR00305	702	881
2075	2076	RXA01092	GR00305	702	881
2077	2078	RXA01096	GR00306	4341	3643
2079	2080	RXA01102	GR00306	10018	8774
2081	2082	RXA01103	GR00306	10316	10092
2083	2084	RXA01107	GR00306	13612	14811
2085	2086	RXA01108	GR00306	15562	14912
2087	2088	RXA01109	GR00306	16281	15640
2089	2090	RXA01119	GR00310	1068	139
2091	2092	RXA01122	GR00311	557	36
2093	2094	RXA01123	GR00311	1090	644
2095	2096	RXA01127	GR00314	2	280
2097	2098	RXA01129	GR00314	1461	3326
2099	2100	RXA01131	GR00315	445	1311
2101	2102	RXA01137	GR00318	1101	1460
2103	2104	RXA01156	GR00327	1588	1388
2105	2106	RXA01158	GR00328	2580	1639
2107	2108	RXA01159	GR00328	3089	2775
2109	2110	RXA01160	GR00328	4187	3213
2111	2112	RXA01163	GR00331	710	6
2113	2114	RXA01165	GR00332	2155	1583
2115	2116	RXA01166	GR00332	3005	2523
2117	2118	RXA01170	GR00334	638	1120
2119	2120	RXA01171	GR00334	1714	2406
2121	2122	RXA01176	GR00335	1980	1477
2123	2124	RXA01177	GR00335	2121	4106
2125	2126	RXA01178	GR00335	4106	4555
2127	2128	RXA01184	GR00338	1489	17
2129	2130	RXA01186	GR00338	3742	2645
2131	2132	RXA01186	GR00338	3742	2645
2133	2134	RXA01187	GR00338	3850	4308
2135	2136	RXA01195	GR00343	1413	1859
2137	2138	RXA01196	GR00343	1889	2578
2139	2140	RXA01197	GR00343	3333	2881
2141	2142	RXA01198	GR00343	3422	3724
2143	2144	RXA01207	GR00347	126	773
2145	2146	RXA01213	GR00351	1508	282
2147	2148	RXA01218	GR00353	1078	1506

2157	2158	RXA01271	GR00367	23467	21656
2159	2160	RXA01273	GR00367	26475	25042
2161	2162	RXA01282	GR00369	5444	4665
2163	2164	RXA01294	GR00373	3537	2872
2165	2166	RXA01295	GR00373	3764	4738
2167	2168	RXA01304	GR00376	1982	2467
2169	2170	RXA01310	GR00380	803	477
2171	2172	RXA01313	GR00381	1116	172
2173	2174	RXA01315	GR00382	1394	744
2175	2176	RXA01316	GR00382	1855	1553
2177	2178	RXA01317	GR00382	2296	1877
2179	2180	RXA01318	GR00382	3616	2315
2181	2182	RXA01330	GR00387	569	1024
2183	2184	RXA01333	GR00389	1231	227
2185	2186	RXA01336	GR00389	3640	3038
2187	2188	RXA01342	GR00389	11296	12807
2189	2190	RXA01348	GR00392	261	752
2191	2192	RXA01349	GR00392	1531	755
2193	2194	RXA01357	GR00393	4357	4659
2195	2196	RXA01359	GR00393	6857	8038
2197	2198	RXA01366	GR00397	1369	980
2199	2200	RXA01367	GR00397	1518	1919
2201	2202	RXA01370	GR00398	1875	2225
2203	2204	RXA01372	GR00399	1	591
2205	2206	RXA01378	GR00401	1281	4
2207	2208	RXA01380	GR00403	2	2017
2209	2210	RXA01384	GR00406	3238	1523
2211	2212	RXA01396	GR00408	6475	6218
2213	2214	RXA01397	GR00408	6894	6475
2215	2216	RXA01401	GR00409	3193	3453
2217	2218	RXA01402	GR00409	3508	3981
2219	2220	RXA01405	GR00410	1844	1389
2221	2222	RXA01413	GR00412	854	1453
2223	2224	RXA01414	GR00412	1628	2134
2225	2226	RXA01417	GR00414	645	49
2227	2228	RXA01421	GR00416	1215	829
2229	2230	RXA01425	GR00417	1701	2585
2231	2232	RXA01429	GR00417	5651	6268
2233	2234	RXA01439	GR00418	5949	6494
2235	2236	RXA01440	GR00418	7496	6489
2237	2238	RXA01441	GR00418	8542	7514
2239	2240	RXA01445	GR00418	15083	14091
2241	2242	RXA01447	GR00418	17885	18733
2243	2244	RXA01452	GR00419	2363	2641
2245	2246	RXA01456	GR00420	898	1419
2247	2248	RXA01457	GR00420	1499	2173
2249	2250	RXA01463	GR00421	2493	1330
2251	2252	RXA01469	GR00422	2091	3122
2253	2254	RXA01470	GR00422	4112	3687
2255	2256	RXA01472	GR00422	5783	5328
2257	2258	RXA01473	GR00422	6596	5832
2259	2260	RXA01474	GR00422	6678	7223
		RXA01475	GR00422	7651	7226

-21-

2271	2272	RXA01501	GR00424	8130	7843
2273	2274	RXA01504	GR00424	10710	11318
2275	2276	RXA01505	GR00424	11318	11815
2277	2278	RXA01506	GR00424	11815	12225
2279	2280	RXA01507	GR00424	12239	12661
2281	2282	RXA01519	GR00424	23725	24471
2283	2284	RXA01520	GR00424	24784	25167
2285	2286	RXA01523	GR00424	27951	28901
2287	2288	RXA01525	GR00424	32301	30580
2289	2290	RXA01527	GR00425	5126	2616
2291	2292	RXA01536	GR00427	4066	2825
2293	2294	RXA01540	GR00428	3083	2382
2295	2296	RXA01543	GR00430	2802	37
2297	2298	RXA01544	GR00430	3496	2897
2299	2300	RXA01545	GR00430	4838	3588
2301	2302	RXA01546	GR00430	5584	4889
2303	2304	RXA01547	GR00430	6371	5709
2305	2306	RXA01548	GR00430	7432	6425
2307	2308	RXA01552	GR00431	6122	5145
2309	2310	RXA01554	GR00432	3719	1578
2311	2312	RXA01560	GR00435	767	438
2313	2314	RXA01575	GR00438	8024	7005
2315	2316	RXA01577	GR00438	8811	9185
2317	2318	RXA01579	GR00439	671	1054
2319	2320	RXA01585	GR00441	1226	600
2321	2322	RXA01586	GR00441	1597	1229
2323	2324	RXA01595	GR00447	3326	4285
2325	2326	RXA01600	GR00447	10460	11128
2327	2328	RXA01602	GR00447	13591	12062
2329	2330	RXA01605	GR00448	960	2474
2331	2332	RXA01610	GR00449	4343	3615
2333	2334	RXA01611	GR00449	4832	4476
2335	2336	RXA01612	GR00449	5235	4891
2337	2338	RXA01619	GR00451	2407	1433
2339	2340	RXA01622	GR00452	1908	2510
2341	2342	RXA01623	GR00452	2514	3224
2343	2344	RXA01624	GR00452	3220	3564
2345	2346	RXA01628	GR00453	866	1879
2347	2348	RXA01630	GR00454	341	1417
2349	2350	RXA01641	GR00456	5182	6552
2351	2352	RXA01642	GR00456	6557	7798
2353	2354	RXA01643	GR00456	8374	7949
2355	2356	RXA01645	GR00456	10574	9969
2357	2358	RXA01646	GR00456	11513	10695
2359	2360	RXA01656	GR00460	1548	2444
2361	2362	RXA01665	GR00463	2152	1433
2363	2364	RXA01671	GR00466	854	1468
2365	2366	RXA01673	GR00467	1807	773
2367	2368	RXA01675	GR00467	2824	3234
2369	2370	RXA01676	GR00467	4179	3424
2371	2372	RXA01677	GR00467	5043	4300
				10681	11313

2385	2386	RXA01729	GR00489	2636	3154
2387	2388	RXA01731	GR00491	109	807
2389	2390	RXA01738	GR00493	3971	4684
2391	2392	RXA01741	GR00493	7535	6738
2393	2394	RXA01748	GR00495	3681	4460
2395	2396	RXA01749	GR00495	4633	6249
2397	2398	RXA01750	GR00496	1878	3518
2399	2400	RXA01752	GR00497	557	6
2401	2402	RXA01753	GR00497	2095	557
2403	2404	RXA01760	GR00498	5095	5376
2405	2406	RXA01768	GR00501	827	450
2407	2408	RXA01770	GR00501	5134	1370
2409	2410	RXA01773	GR00503	34	444
2411	2412	RXA01775	GR00504	178	741
2413	2414	RXA01776	GR00504	838	2289
2415	2416	RXA01777	GR00504	2319	2777
2417	2418	RXA01778	GR00504	2912	4048
2419	2420	RXA01779	GR00504	4246	5664
2421	2422	RXA01780	GR00504	5721	6095
2423	2424	RXA01781	GR00504	6052	6312
2425	2426	RXA01782	GR00504	6384	6779
2427	2428	RXA01783	GR00504	6842	7078
2429	2430	RXA01785	GR00505	729	1304
2431	2432	RXA01788	GR00506	361	801
2433	2434	RXA01789	GR00506	875	1516
2435	2436	RXA01790	GR00506	1672	1731
2437	2438	RXA01791	GR00506	1885	2247
2439	2440	RXA01792	GR00506	2310	2582
2441	2442	RXA01793	GR00506	2916	3149
2443	2444	RXA01794	GR00506	3194	3427
2445	2446	RXA01799	GR00509	377	1570
2447	2448	RXA01800	GR00509	2292	1573
2449	2450	RXA01804	GR00509	6117	5797
2451	2452	RXA01805	GR00509	6515	6186
			GR00509	6595	7074

-23-

2499	2500	RXA01959	GR00564	1639	2019
2501	2502	RXA01961	GR00565	521	1000
2503	2504	RXA01962	GR00565	1022	1591
2505	2506	RXA01963	GR00565	1757	2440
2507	2508	RXA01964	GR00566	1329	4
2509	2510	RXA01965	GR00566	1935	1375
2511	2512	RXA01966	GR00567	47	703
2513	2514	RXA01968	GR00567	3295	2138
2515	2516	RXA01969	GR00567	5689	5216
2517	2518	RXA01973	GR00570	2	583
2519	2520	RXA01974	GR00570	658	2109
2521	2522	RXA01976	GR00571	3742	2222
2523	2524	RXA01977	GR00571	4547	3972
2525	2526	RXA01982	GR00573	3001	1844
2527	2528	RXA01990	GR00581	1	999
2529	2530	RXA01992	GR00583	709	260
2531	2532	RXA01999	GR00589	2384	2854
2533	2534	RXA02001	GR00590	700	152
2535	2536	RXA02004	GR00594	3	209
2537	2538	RXA02006	GR00597	498	4
2539	2540	RXA02009	GR00601	127	5
2541	2542	RXA02013	GR00607	553	5
2543	2544	RXA02021	GR00613	2008	1061
2545	2546	RXA02036	GR00619	3441	3821
2547	2548	RXA02040	GR00621	1452	925
2549	2550	RXA02046	GR00623	2680	2943
2551	2552	RXA02051	GR00624	3186	3683
2553	2554	RXA02053	GR00624	5484	6062
2555	2556	RXA02057	GR00625	2972	3502
2557	2558	RXA02058	GR00625	4051	3500
2559	2560	RXA02069	GR00627	1116	1694
2561	2562	RXA02070	GR00627	1733	2830
2563	2564	RXA02080	GR00628	11017	10211
2565	2566	RXA02081	GR00628	12307	13935
2567	2568	RXA02084	GR00629	2920	2576
2569	2570	RXA02089	GR00629	8431	8901
2571	2572	RXA02090	GR00629	9764	8964
2573	2574	RXA02091	GR00629	10512	9862
2575	2576	RXA02097	GR00630	184	3555
2577	2578	RXA02102	GR00631	4479	3322
2579	2580	RXA02103	GR00631	4510	4905
2581	2582	RXA02109	GR00632	3460	2540
2583	2584	RXA02117	GR00636	1056	1529
2585	2586	RXA02123	GR00636	6558	7928
2587	2588	RXA02124	GR00636	7956	9911
2589	2590	RXA02125	GR00637	739	1539
2591	2592	RXA02129	GR00637	5906	6139
2593	2594	RXA02132	GR00638	737	1375
2595	2596	RXA02137	GR00639	4166	3369
2597	2598	RXA02141	GR00639	8457	8864
2599	2600	RXA02146	GR00639	14742	15368
2601	2602	RXA02152	GR00640	237	638
2603	2604	RXA02163	GR00640	10072	10824
		RXA02164	GR00640	10824	12398

2613	2614	RXA02170	GR00641	4798	4025
2615	2616	RXA02172	GR00641	6919	6581
2617	2618	RXA02177	GR00641	12683	13615
2619	2620	RXA02178	GR00641	13628	14497
2621	2622	RXA02181	GR00641	17168	17845
2623	2624	RXA02183	GR00641	18663	19187
2625	2626	RXA02187	GR00641	21249	23447
2627	2628	RXA02199	GR00646	2591	3160
2629	2630	RXA02203	GR00646	7469	7092
2631	2632	RXA02206	GR00646	9927	10862
2633	2634	RXA02211	GR00648	2537	2989
2635	2636	RXA02212	GR00649	964	467
2637	2638	RXA02216	GR00651	2	307
2639	2640	RXA02217	GR00651	968	306
2641	2642	RXA02218	GR00651	1299	1565
2643	2644	RXA02219	GR00651	1578	2963
2645	2646	RXA02221	GR00651	6720	8081
2647	2648	RXA02227	GR00653	1236	1853
2649	2650	RXA02230	GR00653	4156	3620
2651	2652	RXA02231	GR00653	5111	4356
2653	2654	RXA02244	GR00654	12058	13590
2655	2656	RXA02255	GR00654	22507	23442
2657	2658	RXA02266	GR00655	653	1165
2659	2660	RXA02267	GR00655	2053	1181
2661	2662	RXA02280	GR00658	2	754
2663	2664	RXA02286	GR00660	3285	3833
2665	2666	RXA02287	GR00660	4071	4622
2667	2668	RXA02294	GR00662	5992	5618
2669	2670	RXA02295	GR00662	6842	6063
2671	2672	RXA02297	GR00662	7502	8638
2673	2674	RXA02298	GR00662	10310	8652
2675	2676	RXA02304	GR00663	1613	723
2677	2678	RXA02308	GR00664	939	511
2679	2680	RXA02324	GR00668	1548	2633
2681	2682	RXA02325	GR00668	4314	3445
2683	2684	RXA02331	GR00671	396	761
2685	2686	RXA02336	GR00672	2731	2552
2687	2688	RXA02347	GR00677	509	189
2689	2690	RXA02349	GR00678	394	5
2691	2692	RXA02352	GR00681	2	556
2693	2694	RXA02356	GR00684	761	1756
2695	2696	RXA02358	GR00685	1239	1529
2697	2698	RXA02362	GR00685	7045	10743
2699	2700	RXA02374	GR00688	1626	2246
2701	2702	RXA02390	GR00695	1500	832
2703	2704	RXA02393	GR00697	168	449
2705	2706	RXA02395	GR00698	2	733
2707	2708	RXA02396	GR00698	1309	1031
2709	2710	RXA02403	GR00700	896	1660
2711	2712	RXA02412	GR00703	2043	2522
2713	2714	RXA02417	GR00705	4755	2632
2715	2716	RXA02421	GR00705	7237	6428
2717	2718	RXA02425	GR00707	1	630
2719	2720	RXA02427	GR00707	3447	3061
2721	2722	RXA02430	GR00707	7498	7683
				2004	2580

-25-

2727	2728	RXA02443	GR00709	6818	7771
2729	2730	RXA02444	GR00709	7836	9113
2731	2732	RXA02452	GR00710	5271	5092
2733	2734	RXA02459	GR00712	4341	5075
2735	2736	RXA02461	GR00712	6252	5845
2737	2738	RXA02467	GR00714	643	419
2739	2740	RXA02472	GR00715	5435	5725
2741	2742	RXA02473	GR00715	6664	5924
2743	2744	RXA02475	GR00715	9595	8441
2745	2746	RXA02478	GR00716	1245	10
2747	2748	RXA02482	GR00718	914	105
2749	2750	RXA02483	GR00718	1813	1001
2751	2752	RXA02484	GR00718	2317	1817
2753	2754	RXA02486	GR00718	3441	4076
2755	2756	RXA02488	GR00719	1	369
2757	2758	RXA02489	GR00719	373	996
2759	2760	RXA02495	GR00720	9002	6435
2761	2762	RXA02496	GR00720	10025	9219
2763	2764	RXA02498	GR00720	11016	11819
2765	2766	RXA02500	GR00720	13460	13558
2767	2768	RXA02506	GR00720	19484	18603
2769	2770	RXA02514	GR00723	1	837
2771	2772	RXA02518	GR00723	3464	3874
2773	2774	RXA02521	GR00724	2924	4366
2775	2776	RXA02524	GR00725	2405	3094
2777	2778	RXA02525	GR00725	3113	3490
2779	2780	RXA02540	GR00726	12438	12001
2781	2782	RXA02544	GR00726	16715	18142
2783	2784	RXA02545	GR00726	18749	18192
2785	2786	RXA02546	GR00726	19927	18824
2787	2788	RXA02549	GR00728	1331	6
2789	2790	RXA02552	GR00730	924	130
2791	2792	RXA02554	GR00731	1050	427
2793	2794	RXA02569	GR00736	82	831
2795	2796	RXA02570	GR00736	837	1478
2797	2798	RXA02573	GR00739	594	151
2799	2800	RXA02575	GR00739	1907	3064
2801	2802	RXA02576	GR00740	1569	148
2803	2804	RXA02577	GR00740	2463	1579
2805	2806	RXA02584	GR00741	8925	8575
2807	2808	RXA02585	GR00741	9917	8937
2809	2810	RXA02588	GR00741	13037	12354
2811	2812	RXA02591	GR00741	15780	17609
2813	2814	RXA02598	GR00742	2576	3166
2815	2816	RXA02600	GR00742	5027	3630
2817	2818	RXA02601	GR00742	5258	7246
2819	2820	RXA02602	GR00742	7239	7742
2821	2822	RXA02604	GR00742	8800	10875
2823	2824	RXA02609	GR00742	16197	16445
2825	2826	RXA02617	GR00745	1404	1910
2827	2828	RXA02619	GR00746	204	1103
2829	2830	RXA02620	GR00746	1192	1845
2831	2832	RXA02639	GR00749	511	1344
2833	2834	RXA02647	GR00751	4155	4616
2835	2836	RXA02649	GR00752	1284	283
2837	2838	RXA02652	GR00752	2973	3551
2839	2840	RXA02655	GR00752	9313	8330

2841	2842	RXA02662	GR00753	1461	1724
2843	2844	RXA02665	GR00753	6497	6018
2845	2846	RXA02670	GR00753	10199	10780
2847	2848	RXA02672	GR00753	12303	13400
2849	2850	RXA02678	GR00754	3858	4775
2851	2852	RXA02683	GR00754	7742	7065
2853	2854	RXA02685	GR00754	10058	9402
2855	2856	RXA02688	GR00754	12256	12924
2857	2858	RXA02689	GR00754	13405	13064
2859	2860	RXA02690	GR00754	14502	13405
2861	2862	RXA02700	GR00757	3507	4742
2863	2864	RXA02701	GR00757	4838	6145
2865	2866	RXA02712	GR00758	13067	12273
2867	2868	RXA02714	GR00758	14754	14326
2869	2870	RXA02715	GR00758	15847	15458
2871	2872	RXA02721	GR00759	1373	636
2873	2874	RXA02725	GR00760	1478	867
2875	2876	RXA02727	GR00760	6287	5376
2877	2878	RXA02735	GR00763	777	73
2879	2880	RXA02736	GR00763	1753	797
2881	2882	RXA02751	GR00764	6393	5920
2883	2884	RXA02756	GR00766	3851	2961
2885	2886	RXA02757	GR00766	4475	3930
2887	2888	RXA02765	GR00769	3552	2794
2889	2890	RXA02766	GR00770	986	594
2891	2892	RXA02774	GR00773	3	473
2893	2894	RXA02775	GR00773	744	968
2895	2896	RXA02776	GR00773	1713	1372
2897	2898	RXA02777	GR00773	4626	5732
2899	2900	RXA02778	GR00773	10095	10319
2901	2902	RXA02779	GR00773	10617	10895
2903	2904	RXA02780	GR00773	10954	11280
2905	2906	RXA02783	GR00775	845	1393
2907	2908	RXA02784	GR00775	1751	1936
2909	2910	RXA02786	GR00777	2	808
2911	2912	RXA02789	GR00777	5237	5782
2913	2914	RXA02793	GR00777	9385	8684
2915	2916	RXA02796	GR00778	1648	1100
2917	2918	RXA02798	GR00778	2842	4266
2919	2920	RXA02799	GR00780	182	454
2921	2922	RXA02815	GR00796	3	554
2923	2924	RXA02823	GR00804	275	6
2925	2926	RXA02827	GR00812	428	6
			GR00811	356	15



**TABLE 2: GENES IDENTIFIED FROM GENBANK**

GenBank™ Accession No.	Gene Name	Gene Function	Reference
A09073	ppg	Phosphoenol pyruvate carboxylase	Bachmann, B. et al. "DNA fragment coding for phosphoenolpyruvate carboxylase, recombinant DNA carrying said fragment, strains carrying the recombinant DNA and method for producing L-aminino acids using said strains," Patent: EP 0358940-A 3 03/21/90
A45579, A45581, A45583, A45585 A45587		Threonine dehydratase	Moockel, B. et al. "Production of L-isoleucine by means of recombinant micro-organisms with deregulated threonine dehydratase," Patent: WO 9519442-A 5 07/20/95
AB003132	murC; ftsQ; ftsZ		Kobayashi, M. et al. "Cloning, sequencing, and characterization of the ftsZ gene from corynebacterium bacteria," <i>Biochem. Biophys. Res. Commun.</i> , 236(2):383-388 (1997)
AB015023	murC; ftsQ		Wachi, M. et al. "A murC gene from Corynebacterium bacteria," <i>Appl. Microbiol. Biotechnol.</i> , 51(2):223-228 (1999)
AB018530	ftsR		Kimura, E. et al. "Molecular cloning of a novel gene, ftsR, which rescues the detergent sensitivity of a mutant derived from <i>Brevibacterium lactofermentum</i> ," <i>Biosci. Biotechnol. Biochem.</i> , 60(10):1565-1570 (1996)
AB018531	ftsR1; ftsR2		
AB020624	murI	D-glutamate racemase	
AB023377	tkt	transketolase	
AB024708	glbB; gltD	Glutamine 2-oxoglutarate aminotransferase large and small subunits	
AB025424	acn	aconitase	
AB027714	rep	Replication protein	
AB027715	rep; aad	Replication protein; aminoglycoside adenyltransferase	
AF005242	argC	N-acetylglutamate-5-semialdehyde dehydrogenase	
AF005635	glnA	Glutamine synthetase	
AF030405	hisF	cyclase	
AF030520	argG	Argininosuccinate synthetase	
AF031518	argF	Ornithine carbamoyltransferase	
AF036932	aroD	3-dehydroquinate dehydratase	

GenBank™ Accession No.	Gene Name	Gene Function	Reference
AF038548	pyc	Pyruvate carboxylase	
AF038651	dcIAE; apt; rel	Dipeptide-binding protein; adenine phosphoribosyltransferase; GTP pyrophosphokinase	Wehmeier, L. et al. "The role of the <i>Corynebacterium glutamicum</i> rel gene in (p)ppGpp metabolism," <i>Microbiology</i> , 144:1853-1862 (1998)
AF041436	argR	Arginine repressor	
AF045998	impA	Inositol monophosphate phosphatase	
AF048764	argH	Argininosuccinate lyase	
AF049897	argC; argJ; argB; argD; argF; argR; argG; argH	N-acetylglutamylphosphate reductase; ornithine acetyltransferase; N-acetylglutamate kinase; acetylornithine transaminase; ornithine carbamoyltransferase; arginine repressor; argininosuccinate synthase; argininosuccinate lyase	
AF050109	inhA	Enoyl-acyl carrier protein reductase	
AF050166	hisG	ATP phosphoribosyltransferase	
AF051846	hisA	Phosphoribosylformimino-5-amino-1-phosphoribosyl-4-imidazolecarboxamide isomerase	
AF052652	metA	Homoserine O-acetyltransferase	Park, S. et al. "Isolation and analysis of metA, a methionine biosynthetic gene encoding homoserine acetyltransferase in <i>Corynebacterium glutamicum</i> ," <i>Mol. Cells.</i> , 8(3):286-294 (1998)
AF053071	aroB	Dehydroquinase synthetase	
AF060558	hisH	Glutamine amidotransferase	
AF086704	hisE	Phosphoribosyl-ATP-pyrophosphohydrolase	
AF114233	aroA	5-enolpyruvylshikimate 3-phosphate synthase	
AF116184	panD	L-aspartate-alpha-decarboxylase precursor	Dusch, N. et al. "Expression of the <i>Corynebacterium glutamicum</i> panD gene encoding L-aspartate-alpha-decarboxylase leads to pantothenate overproduction in <i>Escherichia coli</i> ," <i>Appl. Environ. Microbiol.</i> , 65(4)1530-1539 (1999)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
AF124518	aroD; aroE	3-dehydroquinase; shikimate dehydrogenase	
AF124600	aroC; aroK; aroB; pepQ	Chorismate synthase; shikimate kinase; 3-dehydroquinase synthase; putative cytoplasmic peptidase	
AF145897	inhA		
AF145898	inhA		
AJ001436	ectP	Transport of ectoine, glycine betaine, proline	Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes: Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP," <i>J. Bacteriol.</i> , 180(22):6005-6012 (1998)
AJ004934	dapD	Tetrahydrodipicolinate succinylase (incomplete <sup>1</sup> )	Wehrmann, A. et al. "Different modes of diaminopimelate synthesis and their role in cell wall integrity: A study with Corynebacterium glutamicum," <i>J. Bacteriol.</i> , 180(12):3159-3165 (1998)
AJ007732	ppc; secG; amt; ocd; soxA	Phosphoenolpyruvate-carboxylase; ?; high affinity ammonium uptake protein; putative ornithine-cyclodecarboxylase; sarcosine oxidase	
AJ010319	ftsY, glnB, glnD; srp; amtP	Involved in cell division; PII protein; uridylyltransferase (uridylyl-removing enzyme); signal recognition particle; low affinity ammonium uptake protein	Jakoby, M. et al. "Nitrogen regulation in Corynebacterium glutamicum; Isolation of genes involved in biochemical characterization of corresponding proteins," <i>FEMS Microbiol.</i> , 173(2):303-310 (1999)
AJ132968	cat	Chloramphenicol acetyl transferase	
AJ224946	mgo	L-malate: quinone oxidoreductase	Molenaar, D. et al. "Biochemical and genetic characterization of the membrane-associated malate dehydrogenase (acceptor) from Corynebacterium glutamicum," <i>Eur. J. Biochem.</i> , 254(2):395-403 (1998)
AJ238250	ndh	NADH dehydrogenase	
AJ238703	porA	Porin	Lichtinger, T. et al. "Biochemical and biophysical characterization of the cell wall porin of Corynebacterium glutamicum: The channel is formed by a low molecular mass polypeptide," <i>Biochemistry</i> , 37(43):15024-15032 (1998)
D17429		Transposable element IS31831	Vertes, A.A. et al. "Isolation and characterization of IS31831, a transposable element from Corynebacterium glutamicum," <i>Mol. Microbiol.</i> , 11(4):739-746 (1994)

אנו מודים לך על שיתוף הפעולה והתמיכה  
בפרויקט זה. אנחנו מקווים שתיהנה  
מהתהליך.

GenBank™ Accession No.	Gene Name	Gene Function	Reference
D84102	odhA	2-oxoglutarate dehydrogenase	Usuda, Y. et al. "Molecular cloning of the Corynebacterium glutamicum (Brevibacterium lactofermentum AJ12036) odhA gene encoding a novel type of 2-oxoglutarate dehydrogenase," <i>Microbiology</i> , 142:3347-3354 (1996)
E01358	hdh; hk	Homoserine dehydrogenase; homoserine kinase	Katsumata, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 1 10/12/87
E01359		Upstream of the start codon of homoserine kinase gene	Katsumata, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 2 10/12/87
E01375		Tryptophan operon	
E01376	trpL; trpE	Leader peptide; anthranilate synthase	Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87
E01377		Promoter and operator regions of tryptophan operon	Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87
E03937		Biotin-synthase	Hatakeyama, K. et al. "DNA fragment containing gene capable of coding biotin synthetase and its utilization," Patent: JP 1992278088-A 1 10/02/92
E04040		Diamino pelargonic acid aminotransferase	Kohama, K. et al. "Gene coding diaminopelargonic acid aminotransferase and desthiobiotin synthetase and its utilization," Patent: JP 1992330284-A 1 11/18/92
E04041		Desthiobiotinsynthetase	Kohama, K. et al. "Gene coding diaminopelargonic acid aminotransferase and desthiobiotin synthetase and its utilization," Patent: JP 1992330284-A 1 11/18/92
E04307		Flavum aspartase	Kurusu, Y. et al. "Gene DNA coding aspartase and utilization thereof," Patent: JP 1993030977-A 1 02/09/93
E04376		Isocitric acid lyase	Katsumata, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993056782-A 3 03/09/93
E04377		Isocitric acid lyase N-terminal fragment	Katsumata, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993056782-A 3 03/09/93
E04484		Prephenate dehydratase	Sotouchi, N. et al. "Production of L-phenylalanine by fermentation," Patent: JP 1993076352-A 2 03/30/93
E05108		Aspartokinase	Fugono, N. et al. "Gene DNA coding Aspartokinase and its use," Patent: JP 1993184366-A 1 07/27/93
E05112		Dihydro-dipichorinate synthetase	Hatakeyama, K. et al. "Gene DNA coding dihydrodipicolinic acid synthetase and its use," Patent: JP 1993184371-A 1 07/27/93

GenBank™ Accession No.	Gene Name	Gene Function	Reference
E05776		Diaminopimelic acid dehydrogenase	Kobayashi, M. et al. "Gene DNA coding Diaminopimelic acid dehydrogenase and its use," Patent: JP 1993284970-A 1 11/02/93
E05779		Threonine synthase	Kohama, K. et al. "Gene DNA coding threonine synthase and its use," Patent: JP 1993284972-A 1 11/02/93
E06110		Prephenate dehydratase	Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent: JP 1993344881-A 1 12/27/93
E06111		Mutated Prephenate dehydratase	Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent: JP 1993344881-A 1 12/27/93
E06146		Acetohydroxy acid synthetase	Inui, M. et al. "Gene capable of coding Acetohydroxy acid synthetase and its use," Patent: JP 1993344893-A 1 12/27/93
E06825		Aspartokinase	Sugimoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94
E06826		Mutated aspartokinase alpha subunit	Sugimoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94
E06827		Mutated aspartokinase alpha subunit	Sugimoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94
E07701	secY		Honno, N. et al. "Gene DNA participating in integration of membraneous protein to membrane," Patent: JP 1994169780-A 1 06/21/94
E08177		Aspartokinase	Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent: JP 1994261766-A 1 09/20/94
E08178, E08179, E08180, E08181, E08182		Feedback inhibition-released Aspartokinase	Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent: JP 1994261766-A 1 09/20/94
E08232		Acetohydroxy-acid isomeroreductase	Inui, M. et al. "Gene DNA coding acetohydroxy acid isomeroreductase," Patent: JP 1994277067-A 1 10/04/94
E08234	secE		Asai, Y. et al. "Gene DNA coding for translocation machinery of protein," Patent: JP 1994277073-A 1 10/04/94
E08643		FT aminotransferase and desthiobiotin synthetase promoter region	Hatakeyama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent: JP 1995031476-A 1 02/03/95
E08646		Biotin synthetase	Hatakeyama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent: JP 1995031476-A 1 02/03/95

GenBank™ Accession No.	Gene Name	Gene Function	Reference
E08649		Aspartase	Kohama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent: JP 1995031478-A 1 02/03/95
E08900		Dihydrodipicolinate reductase	Madori, M. et al. "DNA fragment containing gene coding Dihydrodipicolinate acid reductase and utilization thereof," Patent: JP 1995075578-A 1 03/20/95
E08901		Diaminopimelic acid decarboxylase	Madori, M. et al. "DNA fragment containing gene coding Diaminopimelic acid decarboxylase and utilization thereof," Patent: JP 1995075579-A 1 03/20/95
E12594		Serine hydroxymethyltransferase	Hatakeyama, K. et al. "Production of L-tryptophan," Patent: JP 1997028391-A 1 02/04/97
E12760, E12759, E12758		transposase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12764		Arginyl-tRNA synthetase; diaminopimelic acid decarboxylase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12767		Dihydrodipicolinic acid synthetase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12770		aspartokinase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12773		Dihydrodipicolinic acid reductase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E13655		Glucose-6-phosphate dehydrogenase	Hatakeyama, K. et al. "Glucose-6-phosphate dehydrogenase and DNA capable of coding the same," Patent: JP 1997224661-A 1 09/02/97
L01508	IlvA	Threonine dehydratase	Moeckel, B. et al. "Functional and structural analysis of the threonine dehydratase of Corynebacterium glutamicum," <i>J. Bacteriol.</i> , 174:8065-8072 (1992)
L07603	EC 4.2.1.15	3-deoxy-D-arabinoheptulosonate-7-phosphate synthase	Chen, C. et al. "The cloning and nucleotide sequence of Corynebacterium glutamicum 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase gene," <i>FEBS Microbiol. Lett.</i> , 107:223-230 (1993)
L09232	IlvB; ilvN; ilvC	Acetohydroxy acid synthase large subunit; Acetohydroxy acid synthase small subunit; Acetohydroxy acid isomeroreductase	Keilhauer, C. et al. "Isoleucine synthesis in Corynebacterium glutamicum: molecular analysis of the ilvB-ilvN-ilvC operon," <i>J. Bacteriol.</i> , 175(17):5595-5603 (1993)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
L18874	PtsM	Phosphoenolpyruvate sugar phosphotransferase	Fouet, A et al. "Bacillus subtilis sucrose-specific enzyme II of the phosphotransferase system: expression in Escherichia coli and homology to enzymes II from enteric bacteria," <i>PNAS USA</i> , 84(24):8773-8777 (1987); Lee, J.K. et al. "Nucleotide sequence of the gene encoding the Corynebacterium glutamicum mannose enzyme II and analyses of the deduced protein sequence," <i>FEMS Microbiol. Lett.</i> , 119(1-2):137-145 (1994)
L27123	aceB	Malate synthase	Lee, H-S. et al. "Molecular characterization of aceB, a gene encoding malate synthase in Corynebacterium glutamicum," <i>J. Microbiol. Biotechnol.</i> , 4(4):256-263 (1994)
L27126		Pyruvate kinase	Jetten, M. S. et al. "Structural and functional analysis of pyruvate kinase from Corynebacterium glutamicum," <i>Appl. Environ. Microbiol.</i> , 60(7):2501-2507 (1994)
L28760	aceA	Isocitrate lyase	
L35906	dtxr	Diphtheria toxin repressor	Oguiza, J.A. et al. "Molecular cloning, DNA sequence analysis, and characterization of the Corynebacterium diphtheriae dtxR from Brevibacterium lactofermentum," <i>J. Bacteriol.</i> , 177(2):465-467 (1995)
M13774		Prephenate dehydratase	Follettie, M.T. et al. "Molecular cloning and nucleotide sequence of the Corynebacterium glutamicum pheA gene," <i>J. Bacteriol.</i> , 167:695-702 (1986)
M16175	5S rRNA		Park, Y-H. et al. "Phylogenetic analysis of the coryneform bacteria by 5S rRNA sequences," <i>J. Bacteriol.</i> , 169:1801-1806 (1987)
M16663	trpE	Anthranilate synthase, 5' end	Sano, K. et al. "Structure and function of the trp operon control regions of Brevibacterium lactofermentum, a glutamic-acid-producing bacterium," <i>Gene</i> , 52:191-200 (1987)
M16664	trpA	Tryptophan synthase, 3' end	Sano, K. et al. "Structure and function of the trp operon control regions of Brevibacterium lactofermentum, a glutamic-acid-producing bacterium," <i>Gene</i> , 52:191-200 (1987)
M25819		Phosphoenolpyruvate carboxylase	O'Regan, M. et al. "Cloning and nucleotide sequence of the Phosphoenolpyruvate carboxylase-coding gene of Corynebacterium glutamicum ATCC13032," <i>Gene</i> , 77(2):237-251 (1989)
M85106		23S rRNA gene insertion sequence	Roller, C. et al. "Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," <i>J. Gen. Microbiol.</i> , 138:1167-1175 (1992)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
M85107, M85108		23S rRNA gene insertion sequence	Roller, C. et al. "Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," <i>J. Gen. Microbiol.</i> , 138:1167-1175 (1992)
M89931	aecD; brnQ; yhbW	Beta C-S lyase; branched-chain amino acid uptake carrier; hypothetical protein yhbW	Rossol, I. et al. "The Corynebacterium glutamicum aecD gene encodes a C-S lyase with alpha, beta-elimination activity that degrades aminoethylcysteine," <i>J. Bacteriol.</i> , 174(9):2968-2977 (1992); Tauch, A. et al. "Isoleucine uptake in Corynebacterium glutamicum ATCC 13032 is directed by the brnQ gene product," <i>Arch. Microbiol.</i> , 169(4):303-312 (1998)
S59299	trp	Leader gene (promoter)	Herry, D.M. et al. "Cloning of the trp gene cluster from a tryptophan-hyperproducing strain of Corynebacterium glutamicum: identification of a mutation in the trp leader sequence," <i>Appl. Environ. Microbiol.</i> , 59(3):791-799 (1993)
U11545	trpD	Anthranilate phosphoribosyltransferase	O'Gara, J.P. and Dunican, L.K. (1994) Complete nucleotide sequence of the Corynebacterium glutamicum ATCC 21850 trpD gene." Thesis, Microbiology Department, University College Galway, Ireland.
U13922	cglIM; cglIR; cglIIR	Putative type II 5-cytosine methyltransferase; putative type II restriction endonuclease; putative type I or type III restriction endonuclease	Schafer, A. et al. "Cloning and characterization of a DNA region encoding a stress-sensitive restriction system from Corynebacterium glutamicum ATCC 13032 and analysis of its role in intergeneric conjugation with Escherichia coli," <i>J. Bacteriol.</i> , 176(23):7309-7319 (1994); Schafer, A. et al. "The Corynebacterium glutamicum cglIM gene encoding a 5-cytosine in an McrBC-deficient Escherichia coli strain," <i>Gene</i> , 203(2):95-101 (1997)
U14965	recA		
U31224	ppx		Ankri, S. et al. "Mutations in the Corynebacterium glutamicum proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)
U31225	proC	L-proline: NADP+ 5-oxidoreductase	Ankri, S. et al. "Mutations in the Corynebacterium glutamicum proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)
U31230	obg; proB; unkdh	?-gamma glutamyl kinase; similar to D-isomer specific 2-hydroxyacid dehydrogenases	Ankri, S. et al. "Mutations in the Corynebacterium glutamicum proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)



GenBank™ Accession No.	Gene Name	Gene Function	Reference
U31281	bioB	Biotin synthase	Serebriiskii, I.G., "Two new members of the bio B superfamily: Cloning, sequencing and expression of bio B genes of <i>Methylobacillus flagellatum</i> and <i>Corynebacterium glutamicum</i> ," <i>Gene</i> , 175:15-22 (1996)
U35023	thtR; accBC	Thiosulfate sulfurtransferase; acyl CoA carboxylase	Jager, W. et al. "A <i>Corynebacterium glutamicum</i> gene encoding a two-domain protein similar to biotin carboxylases and biotin-carboxyl-carrier proteins," <i>Arch. Microbiol.</i> , 166(2):76-82 (1996)
U43535	cmr	Multidrug resistance protein	Jager, W. et al. "A <i>Corynebacterium glutamicum</i> gene conferring multidrug resistance in the heterologous host <i>Escherichia coli</i> ," <i>J. Bacteriol.</i> , 179(7):2449-2451 (1997)
U43536	clpB	Heat shock ATP-binding protein	
U53587	aphA-3	3'-5'-aminoglycoside phosphotransferase	
U89648		<i>Corynebacterium glutamicum</i> unidentified sequence involved in histidine biosynthesis, partial sequence	
X04960	trpA; trpB; trpC; trpD; trpE; trpG; trpL	Tryptophan operon	Matsui, K. et al. "Complete nucleotide and deduced amino acid sequences of the <i>Brevibacterium lactofermentum</i> tryptophan operon," <i>Nucleic Acids Res.</i> , 14(24):10113-10114 (1986)
X07563	lys A	DAP decarboxylase (meso-diaminopimelate decarboxylase, EC 4.1.1.20)	Yeh, P. et al. "Nucleic sequence of the lysA gene of <i>Corynebacterium glutamicum</i> and possible mechanisms for modulation of its expression," <i>Mol. Gen. Genet.</i> , 212(1):112-119 (1988)
X14234	EC 4.1.1.31	Phosphoenolpyruvate carboxylase	Eikmanns, B.J. et al. "The Phosphoenolpyruvate carboxylase gene of <i>Corynebacterium glutamicum</i> : Molecular cloning, nucleotide sequence, and expression," <i>Mol. Gen. Genet.</i> , 218(2):330-339 (1989); Lepiniec, L. et al. "Sorghum Phosphoenolpyruvate carboxylase gene family: structure, function and molecular evolution," <i>Plant. Mol. Biol.</i> , 21 (3):487-502 (1993)
X17313	fda	Fructose-bisphosphate aldolase	Von der Osten, C.H. et al. "Molecular cloning, nucleotide sequence and fine-structural analysis of the <i>Corynebacterium glutamicum</i> fda gene: structural comparison of <i>C. glutamicum</i> fructose-1, 6-bisphosphate aldolase to class I and class II aldolases," <i>Mol. Microbiol.</i> ,
X53993	dapA	L-2, 3-dihydrodipicolinate synthetase (EC 4.2.1.52)	Bonnassie, S. et al. "Nucleic sequence of the dapA gene from <i>Corynebacterium glutamicum</i> ," <i>Nucleic Acids Res.</i> , 18(21):6421 (1990)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X54223		AttB-related site	Cianciotto, N. et al. "DNA sequence homology between att B-related sites of Corynebacterium diphtheriae, Corynebacterium ulcerans, Corynebacterium glutamicum, and the attP site of lambda corynebephage," <i>FEMS. Microbiol. Lett.</i> , 66:299-302 (1990)
X54740	argS; lysA	Arginyl-tRNA synthetase; Diaminopimelate decarboxylase	Marcel, T. et al. "Nucleotide sequence and organization of the upstream region of the Corynebacterium glutamicum lysA gene," <i>Mol. Microbiol.</i> , 4(11):1819-1830 (1990)
X55994	trpL; trpE	Putative leader peptide; anthranilate synthase component I	Heery, D.M. et al. "Nucleotide sequence of the Corynebacterium glutamicum trpE gene," <i>Nucleic Acids Res.</i> , 18(23):7138 (1990)
X56037	thrC	Threonine synthase	Han, K.S. et al. "The molecular structure of the Corynebacterium glutamicum threonine synthase gene," <i>Mol. Microbiol.</i> , 4(10):1693-1702 (1990)
X56075	attB-related site	Attachment site	Cianciotto, N. et al. "DNA sequence homology between att B-related sites of Corynebacterium diphtheriae, Corynebacterium ulcerans, Corynebacterium glutamicum, and the attP site of lambda corynebephage," <i>FEMS. Microbiol. Lett.</i> , 66:299-302 (1990)
X57226	lysC-alpha; lysC-beta; asd	Aspartokinase-alpha subunit; Aspartokinase-beta subunit; aspartate beta semialdehyde dehydrogenase	Kalinowski, J. et al. "Genetic and biochemical analysis of the Aspartokinase from Corynebacterium glutamicum," <i>Mol. Microbiol.</i> , 5(5):1197-1204 (1991); Kalinowski, J. et al. "Aspartokinase genes lysC alpha and lysC beta overlap and are adjacent to the aspartate beta-semialdehyde dehydrogenase gene asd in Corynebacterium glutamicum," <i>Mol. Gen. Genet.</i> , 224(3):317-324 (1990)
X59403	gap;pgk; tpi	Glyceraldehyde-3-phosphate; phosphoglycerate kinase; triosephosphate isomerase	Eikmanns, B.J. "Identification, sequence analysis, and expression of a Corynebacterium glutamicum gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase," <i>J. Bacteriol.</i> , 174(19):6076-6086 (1992)
X59404	gdh	Glutamate dehydrogenase	Bormann, E.R. et al. "Molecular analysis of the Corynebacterium glutamicum gdh gene encoding glutamate dehydrogenase," <i>Mol. Microbiol.</i> , 6(3):317-326 (1992)
X60312	lysI	L-lysine permease	Seep-Feldhaus, A.H. et al. "Molecular analysis of the Corynebacterium glutamicum lysI gene involved in lysine uptake," <i>Mol. Microbiol.</i> , 5(12):2995-3005 (1991)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X66078	cop1	Psl protein	Joliff, G. et al. "Cloning and nucleotide sequence of the csp1 gene encoding Psl, one of the two major secreted proteins of Corynebacterium glutamicum: The deduced N-terminal region of Psl is similar to the Mycobacterium antigen 85 complex," <i>Mol. Microbiol.</i> , 6(16):2349-2362 (1992)
X66112	glt	Citrate synthase	Eikmanns, B.J. et al. "Cloning sequence, expression and transcriptional analysis of the Corynebacterium glutamicum gltA gene encoding citrate synthase," <i>Microbiol.</i> , 140:1817-1828 (1994)
X67737	dapB	Dihydrodipicolinate reductase	
X69103	csp2	Surface layer protein PS2	Peyret, J.L. et al. "Characterization of the cspB gene encoding PS2, an ordered surface-layer protein in Corynebacterium glutamicum," <i>Mol. Microbiol.</i> , 9(1):97-109 (1993)
X69104		IS3 related insertion element	Bonamy, C. et al. "Identification of IS1206, a Corynebacterium glutamicum IS3-related insertion sequence and phylogenetic analysis," <i>Mol. Microbiol.</i> , 14(3):571-581 (1994)
X70959	leuA	Isopropylmalate synthase	Patek, M. et al. "Leucine synthesis in Corynebacterium glutamicum: enzyme activities, structure of leuA, and effect of leuA inactivation on lysine synthesis," <i>Appl. Environ. Microbiol.</i> , 60(1):133-140 (1994)
X71489	icd	Isocitrate dehydrogenase (NADP+)	Eikmanns, B.J. et al. "Cloning sequence analysis, expression, and inactivation of the Corynebacterium glutamicum icd gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme," <i>J. Bacteriol.</i> , 177(3):774-782 (1995)
X72855	GDHA	Glutamate dehydrogenase (NADP+)	
X75083, X70584	mtrA	5-methyltryptophan resistance	Heery, D.M. et al. "A sequence from a tryptophan-hyperproducing strain of Corynebacterium glutamicum encoding resistance to 5-methyltryptophan," <i>Biochem. Biophys. Res. Commun.</i> , 201(3):1255-1262 (1994)
X75085	recA		Fitzpatrick, R. et al. "Construction and characterization of recA mutant strains of Corynebacterium glutamicum and Brevibacterium lactofermentum," <i>Appl. Microbiol. Biotechnol.</i> , 42(4):575-580 (1994)
X75504	aceA; thiX	Partial Isocitrate lyase; ?	Reinscheid, D.J. et al. "Characterization of the isocitrate lyase gene from Corynebacterium glutamicum and biochemical analysis of the enzyme," <i>J. Bacteriol.</i> , 176(12):3474-3483 (1994)
X76875		ATPase beta-subunit	Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," <i>Antonie Van Leeuwenhoek</i> , 64:285-305 (1993)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X77034	tuf	Elongation factor Tu	Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," <i>Antonie Van Leeuwenhoek</i> , 64:285-305 (1993)
X77384	recA		Billman-Jacobe, H. "Nucleotide sequence of a recA gene from <i>Corynebacterium glutamicum</i> ," <i>DNA Seq.</i> , 4(6):403-404 (1994)
X78491	aceB	Malate synthase	Reinscheid, D.J. et al. "Malate synthase from <i>Corynebacterium glutamicum</i> pta-ack operon encoding phosphotransacetylase: sequence analysis," <i>Microbiology</i> , 140:3099-3108 (1994)
X80629	16S rDNA	16S ribosomal RNA	Rainey, F.A. et al. "Phylogenetic analysis of the genera <i>Rhodococcus</i> and <i>Norcardia</i> and evidence for the evolutionary origin of the genus <i>Norcardia</i> from within the radiation of <i>Rhodococcus</i> species," <i>Microbiol.</i> , 141:523-528 (1995)
X81191	gluA; gluB; gluC; gluD	Glutamate uptake system	Kronmeyer, W. et al. "Structure of the gluABCD cluster encoding the glutamate uptake system of <i>Corynebacterium glutamicum</i> ," <i>J. Bacteriol.</i> , 177(5):1152-1158 (1995)
X81379	dapE	Succinyldiaminopimelate desuccinylase	Wehrmann, A. et al. "Analysis of different DNA fragments of <i>Corynebacterium glutamicum</i> complementing dapE of <i>Escherichia coli</i> ," <i>Microbiology</i> , 40:3349-56 (1994)
X82061	16S rDNA	16S ribosomal RNA	Ruimy, R. et al. "Phylogeny of the genus <i>Corynebacterium</i> deduced from analyses of small-subunit ribosomal DNA sequences," <i>Int. J. Syst. Bacteriol.</i> , 45(4):740-746 (1995)
X82928	asd; lysC	Aspartate-semialdehyde dehydrogenase; ?	Serebrijski, I. et al. "Multicopy suppression by asd gene and osmotic stress-dependent complementation by heterologous proA in proA mutants," <i>J. Bacteriol.</i> , 177(24):7255-7260 (1995)
X82929	proA	Gamma-glutamyl phosphate reductase	Serebrijski, I. et al. "Multicopy suppression by asd gene and osmotic stress-dependent complementation by heterologous proA in proA mutants," <i>J. Bacteriol.</i> , 177(24):7255-7260 (1995)
X84257	16S rDNA	16S ribosomal RNA	Pascual, C. et al. "Phylogenetic analysis of the genus <i>Corynebacterium</i> based on 16S rRNA gene sequences," <i>Int. J. Syst. Bacteriol.</i> , 45(4):724-728 (1995)
X85965	aroP; dapE	Aromatic amino acid permease; ?	Wehrmann, A. et al. "Functional analysis of sequences adjacent to dapE of <i>Corynebacterium glutamicum</i> proline reveals the presence of aroP, which encodes the aromatic amino acid transporter," <i>J. Bacteriol.</i> , 177(20):5991-5993 (1995)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X86157	argB; argC; argD; argF; argJ	Acetylglutamate kinase; N-acetyl-gamma-glutamyl-phosphate reductase; acetylornithine aminotransferase; ornithine carbamoyltransferase; glutamate N-acetyltransferase	Sakanyan, V. et al. "Genes and enzymes of the acetyl cycle of arginine biosynthesis in <i>Corynebacterium glutamicum</i> : enzyme evolution in the early steps of the arginine pathway," <i>Microbiology</i> , 142:99-108 (1996)
X89084	pta; ackA	Phosphate acetyltransferase; acetate kinase	Reinscheid, D.J. et al. "Cloning, sequence analysis, expression and inactivation of the <i>Corynebacterium glutamicum</i> pta-ack operon encoding phosphotransacetylase and acetate kinase," <i>Microbiology</i> , 145:503-513 (1999)
X89850	attB	Attachment site	Le Marrec, C. et al. "Genetic characterization of site-specific integration functions of phi AAU2 infecting "Arthrobacter aureus C70," <i>J. Bacteriol.</i> , 178(7):1996-2004 (1996)
X90356		Promoter fragment F1	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90357		Promoter fragment F2	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90358		Promoter fragment F10	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90359		Promoter fragment F13	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90360		Promoter fragment F22	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90361		Promoter fragment F34	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90362		Promoter fragment F37	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X90363		Promoter fragment F45	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90364		Promoter fragment F64	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90365		Promoter fragment F75	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90366		Promoter fragment PF101	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90367		Promoter fragment PF104	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90368		Promoter fragment PF109	Patek, M. et al. "Promoters from Corynebacterium glutamicum: cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X93513	amt	Ammonium transport system	Siewe, R.M. et al. "Functional and genetic characterization of the (methyl) ammonium uptake carrier of Corynebacterium glutamicum," <i>J. Biol. Chem.</i> , 271(10):5398-5403 (1996)
X93514	betP	Glycine betaine transport system	Peter, H. et al. "Isolation, characterization, and expression of the Corynebacterium glutamicum betP gene, encoding the transport system for the compatible solute glycine betaine," <i>J. Bacteriol.</i> , 178(17):5229-5234 (1996)
X95649	orf4		Patek, M. et al. "Identification and transcriptional analysis of the dapB-ORF2-dapA-ORF4 operon of Corynebacterium glutamicum, encoding two enzymes involved in L-lysine synthesis," <i>Biotechnol. Lett.</i> , 19:1113-1117 (1997)
X96471	lysE; lysG	Lysine exporter protein; Lysine export regulator protein	Vrljic, M. et al. "A new type of transporter with a new type of cellular function: L-lysine export from Corynebacterium glutamicum," <i>Mol. Microbiol.</i> , 22(5):815-826 (1996)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X96580	panB; panC; xylB	3-methyl-2-oxobutanoate hydroxymethyltransferase; pantoate-beta-alanine ligase; xylulokinase	Sahm, H. et al. "D-pantothenate synthesis in Corynebacterium glutamicum and use of panBC and genes encoding L-valine synthesis for D-pantothenate overproduction," <i>Appl. Environ. Microbiol.</i> , 65(5):1973-1979 (1999)
X96962		Insertion sequence IS1207 and transposase	
X99289		Elongation factor P	Ramos, A. et al. "Cloning, sequencing and expression of the gene encoding elongation factor P in the amino-acid producer Brevibacterium lactofermentum (Corynebacterium glutamicum ATCC 13869)," <i>Gene</i> , 198:217-222 (1997)
Y00140	thrB	Homoserine kinase	Mateos, L.M. et al. "Nucleotide sequence of the homoserine kinase (thrB) gene of the Brevibacterium lactofermentum," <i>Nucleic Acids Res.</i> , 15(9):3922 (1987)
Y00151	ddh	Meso-diaminopimelate D-dehydrogenase (EC 1.4.1.16)	Ishino, S. et al. "Nucleotide sequence of the meso-diaminopimelate D-dehydrogenase gene from Corynebacterium glutamicum," <i>Nucleic Acids Res.</i> , 15(9):3917 (1987)
Y00476	thrA	Homoserine dehydrogenase	Mateos, L.M. et al. "Nucleotide sequence of the homoserine dehydrogenase (thrA) gene of the Brevibacterium lactofermentum," <i>Nucleic Acids Res.</i> , 15(24):10598 (1987)
Y00546	hom; thrB	Homoserine dehydrogenase; homoserine kinase	Peoples, O.P. et al. "Nucleotide sequence and fine structural analysis of the Corynebacterium glutamicum hom-thrB operon," <i>Mol. Microbiol.</i> , 2(1):63-72 (1988)
Y08964	murC; ftsQ/divD; ftsZ	UPD-N-acetylmuramate-alanine ligase; division initiation protein or cell division protein; cell division protein	Honrubia, M.P. et al. "Identification, characterization, and chromosomal organization of the ftsZ gene from Brevibacterium lactofermentum," <i>Mol. Gen. Genet.</i> , 259(1):97-104 (1998)
Y09163	putP	High affinity proline transport system	Peter, H. et al. "Isolation of the putP gene of Corynebacterium glutamicumproline and characterization of a low-affinity uptake system for compatible solutes," <i>Arch. Microbiol.</i> , 168(2):143-151 (1997)
Y09548	pyc	Pyruvate carboxylase	Peters-Wendisch, P.G. et al. "Pyruvate carboxylase from Corynebacterium glutamicum: characterization, expression and inactivation of the pyc gene," <i>Microbiology</i> , 144:915-927 (1998)
Y09578	leuB	3-isopropylmalate dehydrogenase	Patek, M. et al. "Analysis of the leuB gene from Corynebacterium glutamicum," <i>Appl. Microbiol. Biotechnol.</i> , 50(1):42-47 (1998)
Y12472		Attachment site bacteriophage Phi-16	Moreau, S. et al. "Site-specific integration of corynephage Phi-16: The construction of an integration vector," <i>Microbiol.</i> , 145:539-548 (1999)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
Y12537	proP	Proline/ectoine uptake system protein	Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes: Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP," <i>J. Bacteriol.</i> , 180(22):6005-6012 (1998)
Y13221	glnA	Glutamine synthetase I	Jakoby, M. et al. "Isolation of Corynebacterium glutamicum glnA gene encoding glutamine synthetase I," <i>FEMS Microbiol. Lett.</i> , 154(1):81-88 (1997)
Y16642	lpd	Dihydrolipoamide dehydrogenase	
Y18059		Attachment site Corynebacterium 304L	Moreau, S. et al. "Analysis of the integration functions of &phi;304L: An integrase module among corynephages," <i>Virology</i> , 255(1): 150-159 (1999)
Z21501	argS; lysA	Arginyl-tRNA synthetase; diaminopimelate decarboxylase (partial)	Oguiza, J.A. et al. "A gene encoding arginyl-tRNA synthetase is located in the upstream region of the lysA gene in Brevibacterium lactofermentum: Regulation of argS-lysA cluster expression by arginine," <i>J. Bacteriol.</i> , 175(22):7356-7362 (1993)
Z21502	dapA; dapB	Dihydrodipicolinate synthase; dihydrodipicolinate reductase	Pisabarro, A. et al. "A cluster of three genes (dapA, orf2, and dapB) of Brevibacterium lactofermentum encodes dihydrodipicolinate reductase, and a third polypeptide of unknown function," <i>J. Bacteriol.</i> , 175(9):2743-2749 (1993)
Z29563	thrC	Threonine synthase	Malumbres, M. et al. "Analysis and expression of the thrC gene of the encoded threonine synthase," <i>Appl. Environ. Microbiol.</i> , 60(7):2209-2219 (1994)
Z46753	16S rDNA	Gene for 16S ribosomal RNA	
Z49822	sigA	SigA sigma factor	Oguiza, J.A. et al. "Multiple sigma factor genes in Brevibacterium lactofermentum: Characterization of sigA and sigB," <i>J. Bacteriol.</i> , 178(2):550-553 (1996)
Z49823	galE; dtxR	Catalytic activity UDP-galactose 4-epimerase; diphtheria toxin regulatory protein	Oguiza, J.A. et al. "The galE gene encoding the UDP-galactose 4-epimerase of Brevibacterium lactofermentum is coupled transcriptionally to the dmdR gene," <i>Gene</i> , 177:103-107 (1996)
Z49824	orf1; sigB	?; SigB sigma factor	Oguiza, J.A. et al. "Multiple sigma factor genes in Brevibacterium lactofermentum: Characterization of sigA and sigB," <i>J. Bacteriol.</i> , 178(2):550-553 (1996)
Z66534		Transposase	Correia, A. et al. "Cloning and characterization of an IS-like element present in the genome of Brevibacterium lactofermentum ATCC 13869," <i>Gene</i> , 170(1):91-94 (1996)

<sup>1</sup> A sequence for this gene was published in the indicated reference. However, the sequence obtained by the inventors of the present application is significantly longer than the published version. It is believed that the published version relied on an incorrect start codon, and thus represents only a fragment of the actual coding region.



TABLE 3: Corynebacterium and Brevibacterium Strains Which May be Used in the Practice of the Invention

Genus	Species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NGIC	DSMZ
Brevibacterium	ammoniagenes	21054							
Brevibacterium	ammoniagenes	19350							
Brevibacterium	ammoniagenes	19351							
Brevibacterium	ammoniagenes	19352							
Brevibacterium	ammoniagenes	19353							
Brevibacterium	ammoniagenes	19354							
Brevibacterium	ammoniagenes	19355							
Brevibacterium	ammoniagenes	19356							
Brevibacterium	ammoniagenes	21055							
Brevibacterium	ammoniagenes	21077							
Brevibacterium	ammoniagenes	21553							
Brevibacterium	ammoniagenes	21580							
Brevibacterium	ammoniagenes	39101							
Brevibacterium	butanicum	21196							
Brevibacterium	divaricatum	21792	P928						
Brevibacterium	flavum	21474							
Brevibacterium	flavum	21129							
Brevibacterium	flavum	21518							
Brevibacterium	flavum			B11474					
Brevibacterium	flavum			B11472					
Brevibacterium	flavum	21127							
Brevibacterium	flavum	21128							
Brevibacterium	flavum	21427							
Brevibacterium	flavum	21475							
Brevibacterium	flavum	21517							
Brevibacterium	flavum	21528							
Brevibacterium	flavum	21529							

Brevibacterium	flavum				B11477				
Brevibacterium	flavum				B11478				
Brevibacterium	flavum	21127							
Brevibacterium	flavum				B11474				
Brevibacterium	healii	15527							
Brevibacterium	ketoglutamicum	21004							
Brevibacterium	ketoglutamicum	21089							
Brevibacterium	ketosoreductum	21914							
Brevibacterium	lactofermentum					70			
Brevibacterium	lactofermentum					74			
Brevibacterium	lactofermentum					77			
Brevibacterium	lactofermentum	21798							
Brevibacterium	lactofermentum	21799							
Brevibacterium	lactofermentum	21800							
Brevibacterium	lactofermentum	21801							
Brevibacterium	lactofermentum				B11470				
Brevibacterium	lactofermentum				B11471				
Brevibacterium	lactofermentum	21086							
Brevibacterium	lactofermentum	21420							
Brevibacterium	lactofermentum	21086							
Brevibacterium	lactofermentum	31269							
Brevibacterium	linens	9174							
Brevibacterium	linens	19391							
Brevibacterium	linens	8377							
Brevibacterium	paraffinolyticum					11160			
Brevibacterium	spec.						717.73		
Brevibacterium	spec.						717.73		
Brevibacterium	spec.	14604							
Brevibacterium	spec.	21860							
Brevibacterium	spec.	21864							
Brevibacterium	spec.	21865							

Brevibacterium	spec.	21866						
Brevibacterium	spec.	19240						
Corynebacterium	acetoacidophilum	21476						
Corynebacterium	acetoacidophilum	13870						
Corynebacterium	acetoglutamicum			B11473				
Corynebacterium	acetoglutamicum			B11475				
Corynebacterium	acetoglutamicum	15806						
Corynebacterium	acetoglutamicum	21491						
Corynebacterium	acetoglutamicum	31270						
Corynebacterium	acetophilum			B3671				
Corynebacterium	ammoniagenes	6872						2399
Corynebacterium	ammoniagenes	15511						
Corynebacterium	fujikense	21496						
Corynebacterium	glutamicum	14067						
Corynebacterium	glutamicum	39137						
Corynebacterium	glutamicum	21254						
Corynebacterium	glutamicum	21255						
Corynebacterium	glutamicum	31830						
Corynebacterium	glutamicum	13032						
Corynebacterium	glutamicum	14305						
Corynebacterium	glutamicum	15455						
Corynebacterium	glutamicum	13058						
Corynebacterium	glutamicum	13059						
Corynebacterium	glutamicum	13060						
Corynebacterium	glutamicum	21492						
Corynebacterium	glutamicum	21513						
Corynebacterium	glutamicum	21526						
Corynebacterium	glutamicum	21543						
Corynebacterium	glutamicum	13287						
Corynebacterium	glutamicum	21851						
Corynebacterium	glutamicum	21253						

Corynebacterium	glutamicum	21514							
Corynebacterium	glutamicum	21516							
Corynebacterium	glutamicum	21299							
Corynebacterium	glutamicum	21300							
Corynebacterium	glutamicum	39684							
Corynebacterium	glutamicum	21488							
Corynebacterium	glutamicum	21649							
Corynebacterium	glutamicum	21650							
Corynebacterium	glutamicum	19223							
Corynebacterium	glutamicum	13869							
Corynebacterium	glutamicum	21157							
Corynebacterium	glutamicum	21158							
Corynebacterium	glutamicum	21159							
Corynebacterium	glutamicum	21355							
Corynebacterium	glutamicum	31808							
Corynebacterium	glutamicum	21674							
Corynebacterium	glutamicum	21562							
Corynebacterium	glutamicum	21563							
Corynebacterium	glutamicum	21564							
Corynebacterium	glutamicum	21565							
Corynebacterium	glutamicum	21566							
Corynebacterium	glutamicum	21567							
Corynebacterium	glutamicum	21568							
Corynebacterium	glutamicum	21569							
Corynebacterium	glutamicum	21570							
Corynebacterium	glutamicum	21571							
Corynebacterium	glutamicum	21572							
Corynebacterium	glutamicum	21573							
Corynebacterium	glutamicum	21579							
Corynebacterium	glutamicum	19049							
Corynebacterium	glutamicum	19050							

[illegible]



TABLE 4: ALIGNMENT RESULTS

ID #	length (NT)	Genbank Hit	Length	Accession	Name of Genbank Hit	Source of Genbank Hit	% homology (GAP)	Date of Deposit
rx00003	864	GB_BA2:MPAE000013 GB_BA2:MPAE000013	10328 10328	AE000013 AE000013	Mycoplasma pneumoniae section 13 of 63 of the complete genome. Mycoplasma pneumoniae section 13 of 63 of the complete genome.	Mycoplasma pneumoniae Mycoplasma pneumoniae	37,409 36,768	18-Nov-96 18-Nov-96
rx00008	615	GB_HTG2:AC007356 GB_HTG2:AC007356 GB_HTG2:AC007356 GB_EST36:AV194293	185382 185382 185382 380	AC007356 AC007356 AC007356 AV194293	Drosophila melanogaster chromosome 2 clone BACR24H09 (D595) RPCI-98 24.H.9 map 49A-49B strain Y, cn bw sp. *** SEQUENCING IN PROGRESS ***. 13 unordered pieces. Drosophila melanogaster chromosome 2 clone BACR24H09 (D595) RPCI-98 24.H.9 map 49A-49B strain Y, cn bw sp. *** SEQUENCING IN PROGRESS***. 13 unordered pieces. AV194293 Yuji Kohara unpublished cDNA-Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk627f12 5', mRNA sequence.	Drosophila melanogaster Drosophila melanogaster Drosophila melanogaster Caenorhabditis elegans	39,203 39,203 39,203 38,947	2-Aug-99 2-Aug-99 2-Aug-99 22-Jul-99
rx00015	432	GB_GSS4:AQ684785 GB_PR2:HS217016	671 87552	AQ684785 AL031771	HS_5491_B2_H06_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1057 Col=12 Row=P, genomic survey sequence. Human DNA sequence from clone 217016 on chromosome 6q24 Contains GSS, complete sequence.	Homo sapiens Homo sapiens	41,388 37,471	28-Jun-99 23-Nov-99
rx00018	1422	GB_EST15:AA528550 GB_VI:HS5MCP GB_BA2:AE001270 GB_IN1:LMFL2385 GB_PR4:AC008960 GB_HTG3:AC008266 GB_HTG3:AC008266 GB_EST15:AA496164	335 4320 12448 22004 179757 178972 178972 429	AA528550 M25411 AE001270 AL034389 AC008960 AC008266 AC008266 AA496164	nt01f01.s1 NCI CGAP_Kid1 Homo sapiens cDNA clone IMAGE:912505, mRNA sequence. Human cytomegalovirus major capsid protein (MCP) gene, complete cds. Treponema pallidum section 86 of 87 of the complete genome. Leishmania major Friedlin cosmid L2385, complete sequence. Homo sapiens clone UWGC:dis58 from 7p14-15, complete sequence. Homo sapiens clone DJ1145A24, *** SEQUENCING IN PROGRESS ***. 3 unordered pieces. Homo sapiens clone DJ1145A24, *** SEQUENCING IN PROGRESS ***. 3 unordered pieces. zue67e09.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:743080 5', mRNA sequence.	Homo sapiens human herpesvirus 5 Treponema pallidum Leishmania major Homo sapiens Homo sapiens Homo sapiens Homo sapiens	40,789 38,231 37,130 37,518 36,618 35,419 35,419 35,526	19-Aug-97 30-OCT-1994 16-Jul-98 15-MAR-1999 05-MAR-1999 21-Aug-99 21-Aug-99 11-Aug-97
rx00021	1896	GB_EST30:AI660039 GB_EST37:AI953059 GB_EST15:AA496164	443 522 429	AI660039 AI953059 AA496164	we65d06.x1 Soares_thymus_NHFTn Homo sapiens cDNA clone IMAGE:2345963 3', mRNA sequence. wq49g06.x1 NCI CGAP_GC6 Homo sapiens cDNA clone IMAGE:2474650 3', mRNA sequence. zue67e09.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:743080 5', mRNA sequence.	Homo sapiens Homo sapiens Homo sapiens	42,574 39,198 41,141	10-MAY-1999 6-Sep-99 11-Aug-97
rx00025	1560	GB_PR3:AF022141 GB_EST18:AA678649 GB_PL2:AC009978	43473 538 97554	AF022141 AA678649 AC009978	Homo sapiens chromosome 21q22.2 cosmid Q13F10, complete sequence. ah07c05.s1 Gessler Wilms tumor Homo sapiens cDNA clone IMAGE:1155944 3' similar to gb:X16869 ELONGATION FACTOR 1-ALPHA 1 (HUMAN); mRNA sequence. Genomic sequence for Arabidopsis thaliana BAC T23E18 from chromosome 1, complete sequence.	Homo sapiens Homo sapiens Arabidopsis thaliana	37,262 38,104 34,173	21-Jan-98 02-DEC-1997 15-Nov-99
rx00027	489	GB_HTG2:AC005958 GB_HTG2:AC005958 GB_PR3:HSDJ247C2 GB_PR3:HSDJ247C2	216706 216706 98358 98358	AC005958 AC005958 AL049713 AL049713	Homo sapiens, *** SEQUENCING IN PROGRESS ***. 40 unordered pieces. Homo sapiens, *** SEQUENCING IN PROGRESS ***. 40 unordered pieces. Human DNA sequence from clone 247C2 on chromosome 11p13, complete sequence. Human DNA sequence from clone 247C2 on chromosome 11p13, complete sequence.	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	35,374 35,374 33,056 37,988	11-Nov-98 11-Nov-98 23-Nov-99 23-Nov-99
rx00028								
rx00031	525	GB_PL1:SPBC725 GB_EST5:N22565	37949 435	AL034352 N22565	S.pombe chromosome II cosmid c725. yw30f05.s1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:253761 3', mRNA sequence.	Schizosaccharomyces pombe Homo sapiens	36,084 41,570	29-MAR-1999 20-DEC-1995
rx00049	810	GB_HTG2:HSJ749H19 GB_EST21:AA993042	253387 464	AL117380 AA993042	o192f07.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:1624261 3', mRNA sequence. Homo sapiens chromosome 20 clone RP4-749H19 map q13.11-13.33, *** SEQUENCING IN PROGRESS ***. in unordered pieces.	Homo sapiens Homo sapiens	41,499 37,132	27-Aug-98 03-DEC-1999

TABLE 4: ALIGNMENT RESULTS

GB_HTG2:HSJ749H19	253387	AL117380	Homo sapiens chromosome 20 clone RP4-749H19 map q13.11-13.33, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	37,132	03-DEC-1999
GB_HTG4:AC010137	155817	AC010137	Homo sapiens clone NH0169D01, *** SEQUENCING IN PROGRESS *** 4 unordered pieces.	Homo sapiens	40,052	17-OCT-1999
GB_EST37:AI962012	382	AI962012	w41e06.x1 NCI CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2510050 3' similar to SW:ALC2_HUMAN P01877 IG ALPHA-2 CHAIN C REGION ; mRNA sequence.	Homo sapiens	40,486	20-Aug-99
GB_GSS13:AQ454792	450	AQ454792	HS_5195_B2_H04_SP6E RPCL1-11 Human Male BAC Library Homo sapiens genomic clone Plate=771 Col=8 Row=P, genomic survey sequence.	Homo sapiens	40,991	21-Apr-99
GB_GSS13:AQ454792	450	AQ454792	HS_5195_B2_H04_SP6E RPCL1-11 Human Male BAC Library Homo sapiens genomic clone Plate=771 Col=8 Row=P, genomic survey sequence.	Homo sapiens	40,278	21-Apr-99
GB_GSS5:AAQ773786	459	AQ773786	HS_2222_A1_E07_MR.CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2222 Col=13 Row=I, genomic survey sequence.	Homo sapiens	40,087	29-Jul-99
GB_GSS5:AAQ773786	459	AQ773786	HS_2222_A1_E07_MR.CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2222 Col=13 Row=I, genomic survey sequence.	Homo sapiens	40,087	29-Jul-99
GB_IN1:CEF56H9	28291	Z74473	Caenorhabditis elegans cosmid F56H9, complete sequence.	Caenorhabditis elegans	35,301	23-Nov-98
GB_IN1:CEF56H9	28291	Z74473	Caenorhabditis elegans cosmid F56H9, complete sequence.	Caenorhabditis elegans	38,941	23-Nov-98
GB_HTG6:AC011647	141830	AC011647	Homo sapiens clone RP11-15D18, *** SEQUENCING IN PROGRESS *** 29 unordered pieces.	Homo sapiens	39,939	04-DEC-1999
GB_HTG6:AC011647	141830	AC011647	Homo sapiens clone RP11-15D18, *** SEQUENCING IN PROGRESS *** 29 unordered pieces.	Homo sapiens	37,537	04-DEC-1999
GB_GSS6:AQ825754	463	AQ825754	HS_5441_A2_G02_SP6E RPCL1-11 Human Male BAC Library Homo sapiens genomic clone Plate=1017 Col=4 Row=M, genomic survey sequence.	Homo sapiens	34,444	27-Aug-99
GB_PAT:132939	30001	I32939	Sequence 1 from patent US 5589385.	Unknown.	42,049	6-Feb-97
GB_PAT:AR031772	30001	AR031772	Sequence 1 from patent US 5866410.	Unknown.	42,049	29-Sep-99
GB_PAT:E16763	2517	E16763	gDNA encoding aspartate transferase (AAT).	Corynebacterium glutamicum	98,765	28-Jul-99
GB_EST32:AU050556	813	AU050556	AU050556 Paralicthys olivaceus library (Aoki T) Paralicthys olivaceus cDNA clone WF7-19, mRNA sequence.	Paralicthys olivaceus	35,638	8-Jun-99
GB_EST32:AU050215	733	AU050215	AU050215 Paralicthys olivaceus library (Aoki T) Paralicthys olivaceus cDNA clone WB1-12, mRNA sequence.	Paralicthys olivaceus	35,638	8-Jun-99
GB_HTG3:AC008289	115120	AC008289	Drosophila melanogaster chromosome 2 clone BACR04E05 (D1055) RPCL-98 04.E.5 map 57B-57B strain Y, cn bw sp, *** SEQUENCING IN PROGRESS *** 100 unordered pieces.	Drosophila melanogaster	30,397	17-Aug-99
GB_IN2:AC004433	85862	AC004433	Drosophila melanogaster, chromosome 2R, region 57B1-57B6, P1 clone DS03659, complete sequence.	Drosophila melanogaster	35,501	01-DEC-1998
GB_HTG3:AC008289	115120	AC008289	Drosophila melanogaster chromosome 2 clone BACR04E05 (D1055) RPCL-98 04.E.5 map 57B-57B strain Y, cn bw sp, *** SEQUENCING IN PROGRESS *** 100 unordered pieces.	Drosophila melanogaster	30,397	17-Aug-99
GB_PL2:ATAC006201	87947	AC006201	Arabidopsis thaliana chromosome II BAC T27K22 genomic sequence, complete sequence.	Arabidopsis thaliana	39,099	12-MAR-1999
GB_HTG5:AC010146	271437	AC010146	Homo sapiens clone NH0355113, WORKING DRAFT SEQUENCE, 1 unordered pieces.	Homo sapiens	34,237	12-Nov-99
GB_GSS3:B85079	307	B85079	RPCL11-2909_TP RPCL1-11 Homo sapiens genomic clone RPCL1-11-2909, genomic survey sequence.	Homo sapiens	39,560	9-Apr-99
GB_PR4:AC007157	152937	AC007157	Homo sapiens, clone hRPK 78_A_1, complete sequence.	Homo sapiens	37,661	27-Apr-99
GB_HTG1:CEY43C5	149571	AL021449	Caenorhabditis elegans chromosome IV clone Y43C5, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Caenorhabditis elegans	25,242	23-Jan-98
GB_HTG1:CEY43C5	149571	AL021449	Caenorhabditis elegans chromosome IV clone Y43C5, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Caenorhabditis elegans	38,258	23-Jan-98
GB_IN1:CTAJ2763	1087	AJ002763	Chironomus tentans mRNA for P23 protein (23 kDa).	Chironomus tentans	36,176	26-Jan-98
GB_IN1:CTHRP23	752	AJ003820	Chironomus tentans mRNA for hnRNP protein, hnp23.	Chironomus tentans	36,176	02-DEC-1998
GB_EST17:AA650674	540	AA650674	30788 Lambda-PRL2 Arabidopsis thaliana cDNA clone 277G777, mRNA sequence.	Arabidopsis thaliana	36,965	31-OCT-1997
GB_EST38:AW039986	564	AW039986	EST282477 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA clone cLET19F23, mRNA sequence.	Lycopersicon esculentum	38,078	18-OCT-1999



TABLE 4: ALIGNMENT RESULTS

rx000082	687	GB_EST33:AI778332	378	AI778332	EST259211 tomato susceptible, Cornell Lycopersicon esculentum cDNA clone cLES5A13, mRNA sequence.	Lycopersicon esculentum	38,298	29-Jun-99
		GB_EST38:AW039988	564	AW039988	EST282479 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA clone cLET19F23, mRNA sequence.	Lycopersicon esculentum	38,078	18-OCT-1999
		GB_PR1:HSS171	333303	AJ011930	Homo sapiens chromosome 21q22.3, PAC clones 314N7, 225L15, BAC clone 7B7, complete sequence bases 1...333303.	Homo sapiens	36,111	10-Nov-98
		GB_PR1:HSS171	333303	AJ011930	Homo sapiens chromosome 21q22.3, PAC clones 314N7, 225L15, BAC clone 7B7, complete sequence bases 1...333303.	Homo sapiens	35,432	10-Nov-98
		GB_PR3:AC004000	128117	AC004000	Human PAC clone DJ404F18 from Xq23, complete sequence.	Homo sapiens	38,750	15-Jan-98
rx000083	423	GB_HTG1:CNS0180Y	168868	AL109769	Homo sapiens chromosome 14 clone R-501E21, *** SEQUENCING IN PROGRESS *** , in ordered pieces.	Homo sapiens	33,806	15-OCT-1999
		GB_HTG1:CNS0180Y	168868	AL109769	Homo sapiens chromosome 14 clone R-501E21, *** SEQUENCING IN PROGRESS *** , in ordered pieces.	Homo sapiens	33,806	15-OCT-1999
		GB_PR3:HS516C23	116685	Z93021	Human DNA sequence from clone 516C23 on chromosome 6q12 Contains CA repeat (D6S402) and GSSs, complete sequence.	Homo sapiens	36,562	23-Nov-99
rx000087	651	GB_BA1:PSEBPHABC	6780	M83673	P.pseudoaicaligenes dioxygenase (bphABC) gene cluster, complete cds.	Pseudomonas pseudoaicaligenes	39,564	26-Apr-93
		GB_BA1:PSEBPHA	5700	M86348	Pseudomonas sp. LB400 biphenyl dioxygenase (bphA), biphenyl dioxygenase (bphF) and biphenyl dioxygenase (bphG)s, complete cds, and dihydrodiol dehydrogenase (bphB), partial cds.	Burkholderia sp. LB400	39,564	18-Jul-97
		GB_HTG2:AC007361	4721	E04215	Benzene dioxygenase gene.	Pseudomonas aeruginosa	45,814	29-Sep-97
rx000093	2346	GB_HTG2:AC007361	36465	AC007361	Homo sapiens clone NH0144P23, *** SEQUENCING IN PROGRESS *** , 1 unordered pieces.	Homo sapiens	37,179	23-Apr-99
		GB_PR2:AC006043	189036	AC006043	Homo sapiens BAC clone NH0538D15 from 7q11.23-q21.1, complete sequence.	Homo sapiens	37,060	20-Feb-99
		GB_HTG2:AC007361	36465	AC007361	Homo sapiens clone NH0144P23, *** SEQUENCING IN PROGRESS *** , 1 unordered pieces.	Homo sapiens	37,179	23-Apr-99
		GB_EST15:AA533064	534	AA533064	ri60d06.s1 NCI CGAP_P9 Homo sapiens cDNA clone IMAGE:998875, mRNA sequence.	Homo sapiens	39,024	21-Aug-97
rx000096	426	GB_IN1:CELF01G12	34671	U53342	Caenorhabditis elegans cosmid F01G12.	Caenorhabditis elegans	38,060	5-Apr-96
		GB_PR3:AC004511	45005	AC004511	Homo sapiens chromosome 5, P1 clone 792C12 (LBNL H22), complete sequence.	Homo sapiens	39,163	31-MAR-1998
rx000097	1299	GB_OM:CFU60590	6726	U60590	Canis familiaris TTX-resistant sodium channel mRNA, complete cds.	Canis familiaris	39,528	8-Jan-98
		GB_GSS15:AQ664394	485	AQ664394	HS_5480_B1_B02_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=1056 Col=3 Row=D, genomic survey sequence.	Homo sapiens	39,666	23-Jun-99
		GB_BA2:RSAF000233	5984	AF000233	Rhodobacter sphaeroides nitric oxide reductase operon: norC, norB, norQ, norD, nmrT and nmrU genes, complete cds.	Rhodobacter sphaeroides	37,500	6-Jun-97
rx000101								
rx000108	643	GB_PR4:AC007115	180821	AC007115	Homo sapiens chromosome 12 clone 917O5, complete sequence.	Homo sapiens	35,165	17-Aug-99
		GB_PR3:AC004080	129354	AC004080	Homo sapiens PAC clone DJ0170019 from 7p15-p21, complete sequence.	Homo sapiens	38,560	29-Jan-98
		GB_HTG1:HSJA9613	45302	AJ009613	Homo sapiens chromosome 17 clone cosmid 5L5 map p11, *** SEQUENCING IN PROGRESS *** , in unordered pieces.	Homo sapiens	39,274	11-Nov-98
rx000110	672	GB_PL1:MIATGENB	166924	Y08502	A.thaliana mitochondrial genome, part B.	Mitochondrion Arabidopsis thaliana	36,391	13-Nov-98
		GB_PL2:AC010718	87684	AC010718	Arabidopsis thaliana chromosome I BAC F28O16 genomic sequence, complete sequence.	Arabidopsis thaliana	36,622	30-OCT-1999
		GB_PL2:AC007729	106639	AC007729	Arabidopsis thaliana chromosome II BAC T18C6 genomic sequence, complete sequence.	Arabidopsis thaliana	35,053	5-Jun-99
rx000114	612	GB_OM:BTMICS01	362	Z27071	B.taurus (cos IE3) microsatellite DNA (362bp).	Bos taurus	37,117	10-Aug-95
		GB_OM:BTMICS01	362	Z27071	B.taurus (cos IE3) microsatellite DNA (362bp).	Bos taurus	36,486	10-Aug-95
rx000117	714	GB_PL2:AF080249	3194	AF080249	Arabidopsis thaliana kinesin-like heavy chain (KATD) mRNA, complete cds.	Arabidopsis thaliana	37,846	14-Apr-99
		GB_PL2:IG002P16	110946	AF002720	Arabidopsis thaliana BAC IG002P16.	Arabidopsis thaliana	37,110	12-Jun-97
		GB_PL2:AF080249	3194	AF080249	Arabidopsis thaliana kinesin-like heavy chain (KATD) mRNA, complete cds.	Arabidopsis thaliana	36,506	14-Apr-99
rx000118	378	GB_HTG2:AC008043	124844	AC008043	Drosophila melanogaster chromosome 3 clone BACR05A08 (D750) RPCL-98 05 A.8 map 94A-94A strain y, cn bw sp, *** SEQUENCING IN PROGRESS *** , 86 unordered pieces.	Drosophila melanogaster	33,780	2-Aug-99

TABLE 4: ALIGNMENT RESULTS

GB_HTG2:AC008043	124844	AC008043	Drosophila melanogaster chromosome 3 clone BACR05A08 (D750) RPCI-98 05.A.8 map 94A-94A strain y, cn bw sp *** SEQUENCING IN PROGRESS***, 86 unordered pieces.	Drosophila melanogaster	33,780	2-Aug-99
GB_PR3:AC004827	129690	AC004827	Homo sapiens PAC clone DJ04L15 from Xq23, complete sequence.	Homo sapiens	32,320	17-OCT-1998
GB_PR4:HSU34879	46610	U34879	Human 17-beta-hydroxysteroid dehydrogenase (EDH17B2) gene, complete cds.	Homo sapiens	36,671	14-Jan-99
GB_PR4:HSU34879	46610	U34879	Human 17-beta-hydroxysteroid dehydrogenase (EDH17B2) gene, complete cds.	Homo sapiens	38,345	14-Jan-99
GB_EST37:AW005997	702	AW005997	wz91c01.x1 NCL_CGAP_Bn25 Homo sapiens cDNA clone IMAGE:2566176 3' similar to TR:008609 O08609 TRANSCRIPTION FACTOR-LIKE PROTEIN 4.; mRNA sequence.	Homo sapiens	40,774	10-Sep-99
GB_BA1:TRU80216	1936	U80216	Thermomicrobium roseum 70 kDa heat shock protein Hsp70 (DnaK) gene, complete cds.	Thermomicrobium roseum	38,000	1-Feb-97
GB_HTG1:HS791K14	155318	AL035685	Homo sapiens chromosome 20 clone RP4-791K14, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	37,277	23-Nov-99
GB_HTG1:HS791K14	155318	AL035685	Homo sapiens chromosome 20 clone RP4-791K14, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	37,277	23-Nov-99
GB_HTG1:HSBA29806	198847	AL118525	Homo sapiens chromosome 20 clone RP11-29806, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	36,199	24-Nov-99
GB_HTG1:HS791K14	155318	AL035685	Homo sapiens chromosome 20 clone RP4-791K14, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	36,983	23-Nov-99
GB_HTG1:HS791K14	155318	AL035685	Homo sapiens chromosome 20 clone RP4-791K14, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	36,983	23-Nov-99
GB_PL1:MZECPN60A	6575	L21007	Corn nuclear-encoded mitochondrial chaperonin 60 (cpn60) gene, complete cds.	Zea mays	36,098	26-Jul-93
GB_PL1:ZMCPNAGA	2247	Z12114	Z. mays CPNA gene encoding mitochondrial chaperonin-60.	Zea mays	37,702	01-OCT-1992
GB_PL1:ZMCHHSP60	2138	Z11546	Z. mays mRNA for mitochondrial chaperonin hsp60.	Zea mays	37,721	11-Jun-92
GB_PR4:AC005193	108400	AC005193	Homo sapiens clone DJ065N24, complete sequence.	Homo sapiens	37,500	1-Jul-99
GB_PR4:AC005193	108400	AC005193	Homo sapiens clone DJ065N24, complete sequence.	Homo sapiens	36,796	1-Jul-99
GB_GSS11:AQ299024	449	AQ299024	HS_3178_B1_B06_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3178 Col=11 Row=D, genomic survey sequence.	Homo sapiens	40,757	15-DEC-1998
GB_GSS11:AQ299024	449	AQ299024	HS_3178_B1_B06_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3178 Col=11 Row=D, genomic survey sequence.	Homo sapiens	40,443	15-DEC-1998
GB_GSS10:AQ177172	393	AQ177172	HS_3225_A2_E10_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3225 Col=20 Row=I, genomic survey sequence.	Homo sapiens	50,000	17-OCT-1998
GB_PR3:AC005726	185215	AC005726	Homo sapiens chromosome 17, clone hRPK.192_H_23, complete sequence.	Homo sapiens	37,778	30-OCT-1998
GB_PR3:AC005726	185215	AC005726	Homo sapiens chromosome 17, clone hRPK.192_H_23, complete sequence.	Homo sapiens	38,710	30-OCT-1998
GB_VI:OPU75930	131993	U75930	Orgyia pseudotsugata nuclear polyhedrosis virus complete genome.	Orgyia pseudotsugata nuclear polyhedrosis virus	39,007	06-MAR-1998
GB_HTG2:AC006319	156299	AC006319	Homo sapiens clone DJ0837C09, *** SEQUENCING IN PROGRESS *** 1 unordered pieces.	Homo sapiens	31,773	23-Apr-99
GB_HTG2:AC006319	156299	AC006319	Homo sapiens clone DJ0837C09, *** SEQUENCING IN PROGRESS *** 1 unordered pieces.	Homo sapiens	31,773	23-Apr-99
GB_VI:OPU75930	131993	U75930	Orgyia pseudotsugata nuclear polyhedrosis virus complete genome.	Orgyia pseudotsugata nuclear polyhedrosis virus	38,079	06-MAR-1998
GB_PR4:AC004526	297898	AC004526	Homo sapiens chromosome 17, Neurofibromatosis 1 locus, complete sequence.	Homo sapiens	37,336	25-Feb-99
GB_PR2:HUMNEUROF	100849	L05367	Human oligodendrocyte myelin glycoprotein (OMG) exons 1-2; neurofibromatosis 1 (NF1) exons 28-49; ecotropic viral integration site 2B (EVI2B) exons 1-2; ecotropic viral integration site 2A (EVI2A) exons 1-2; adenylate kinase (AK3) exons 1-2.	Homo sapiens	37,336	20-Sep-95
GB_PR4:HUAC002331	139480	AC002331	Homo sapiens Chromosome 16 BAC clone CIT987SK-A-A-218C7, complete sequence.	Homo sapiens	38,898	23-Nov-99
GB_PR3:AF064861	133965	AF064861	Homo sapiens PAC 128M19 derived from chromosome 21q22.3, containing the HMG-14 and CHD5 genes, complete cds, complete sequence.	Homo sapiens	37,182	2-Jun-98
GB_HTG3:AC009451	165302	AC009451	Homo sapiens chromosome 17 clone 2286_H_12 map 17, *** SEQUENCING IN PROGRESS *** 26 unordered pieces.	Homo sapiens	33,167	22-Aug-99
GB_PR4:AF130343	292721	AF130343	Homo sapiens chromosome 8 clone PAC 87.2 map 8q24.1, complete sequence.	Homo sapiens	36,032	9-Jul-99
GB_HTG4:AC008578	98891	AC008578	Homo sapiens chromosome 5 clone CIT-HSPC_558D4, *** SEQUENCING IN PROGRESS *** 143 unordered pieces.	Homo sapiens	38,129	31-OCT-1999

TABLE 4: ALIGNMENT RESULTS

rx00151	720	GB_PL2:AF058914	111767	AF058914	AC008578	98891	GB_HTG4:AC008578	98891	AC008578	Homo sapiens chromosome 5 clone CIT-HSPC_588D4, *** SEQUENCING IN PROGRESS ***; 143 Homo sapiens unordered pieces.	38,129	31-OCT-1999
		GB_PL2:AF058914	111767	AF058914	AC008578	98891	GB_HTG4:AC008578	98891	AC008578	Arabis thaliana BAC F21E10.	36,068	15-Apr-98
		GB_PL1:AB019440	200000	AB019440	AC000098	100000	GB_PL1:AB019440	200000	AB019440	Homo sapiens DNA for immunoglobulin heavy-chain variable region, complete sequence, 4 of 5.	36,517	24-Feb-99
		GB_PL2:AP000098	100000	AP000098	AC000098	100000	GB_PL2:AP000098	100000	AP000098	Homo sapiens genomic DNA of 21q22.1, GART and AML related, Q78C10-149C3 region, segment 1/20, complete sequence.	39,224	25-Sep-99
rx00153	549	GB_PR4:AC006265	177707	AC006265	AC007389	207188	GB_PR4:AC006265	177707	AC006265	Homo sapiens chromosome 17, clone hRPK_566_B_16, complete sequence.	34,862	28-Jan-99
		GB_HTG2:AC007389	207188	AC007389	AC007389	207188	GB_HTG2:AC007389	207188	AC007389	Homo sapiens clone NH0418H16, *** SEQUENCING IN PROGRESS ***; 6 unordered pieces.	36,044	5-Jun-99
		GB_HTG2:AC007389	207188	AC007389	AC007389	207188	GB_HTG2:AC007389	207188	AC007389	Homo sapiens clone NH0418H16, *** SEQUENCING IN PROGRESS ***; 6 unordered pieces.	36,044	5-Jun-99
rx00154												
rx00155	906	GB_BA2:AE001707	19518	AE001707	AE001707	19518	GB_BA2:AE001707	19518	AE001707	Thermotoga maritima section 19 of 136 of the complete genome.	36,854	2-Jun-99
		GB_PR2:HS112814	19544	AL078589	AL078589	19544	GB_PR2:HS112814	19544	AL078589	Human DNA sequence from clone 112814 on chromosome 6q16.1-16.3. Contains an STS and GSSs, complete sequence.	36,723	23-Nov-99
		GB_BA1:MTCY01B2	35938	Z95554	Z95554	35938	GB_BA1:MTCY01B2	35938	Z95554	Mycobacterium tuberculosis H37Rv complete genome; segment 72/162.	40,435	17-Jun-98
rx00159	1305	GB_EST38:AW048718	475	AW048718	AW048718	475	GB_EST38:AW048718	475	AW048718	UI-M-BH1-amy-d-01-0-UI.s1 NIH_BMAP_M_S2 Mus musculus cDNA clone UI-M-BH1-amy-d-01-0-UI 3', mRNA sequence.	39,789	18-Sep-99
		GB_EST21:AA993450	381	AA993450	AA993450	381	GB_EST21:AA993450	381	AA993450	0132h09.s1 Soares, NHT Homo sapiens cDNA clone IMAGE:1618529 3', mRNA sequence.	38,684	27-Aug-98
rx00161	585	GB_BA1:AB007009	363	AB007009	AB007009	363	GB_BA1:AB007009	363	AB007009	Cytophaga sp. 16S rRNA gene, partial sequence.	39,039	13-OCT-1997
		GB_HTG3:AC009708	25123	AC009708	AC009708	25123	GB_HTG3:AC009708	25123	AC009708	Homo sapiens chromosome 8 clone 318_G_5 map 8, LOW-PASS SEQUENCE SAMPLING.	37,108	28-Aug-99
		GB_HTG3:AC009708	25123	AC009708	AC009708	25123	GB_HTG3:AC009708	25123	AC009708	Homo sapiens chromosome 8 clone 318_G_5 map 8, LOW-PASS SEQUENCE SAMPLING.	37,108	28-Aug-99
		GB_PR3:HSN104C4	40203	Z83855	Z83855	40203	GB_PR3:HSN104C4	40203	Z83855	Human DNA sequence from clone N104C4 on chromosome 22 Contains GSSs, complete sequence.	37,634	23-Nov-99
rx00162	477	GB_HTG1:CEY94A7	41009	Z99294	Z99294	41009	GB_HTG1:CEY94A7	41009	Z99294	Caenorhabditis elegans chromosome V clone Y94A7, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	41,502	18-Sep-97
		GB_HTG1:CEY94A7	41009	Z99294	Z99294	41009	GB_HTG1:CEY94A7	41009	Z99294	Caenorhabditis elegans chromosome V clone Y94A7, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	41,502	18-Sep-97
rx00167	621	GB_BA2:AE001182	11228	AE001182	AE001182	11228	GB_BA2:AE001182	11228	AE001182	Borrelia burgdorferi (section 68 of 70) of the complete genome.	39,655	15-DEC-1997
		GB_HTG7:AC007937	206285	AC007937	AC007937	206285	GB_HTG7:AC007937	206285	AC007937	Mus musculus chromosome 10 clone RP21-538F4 map 10, *** SEQUENCING IN PROGRESS ***; 6 unordered pieces.	37,092	09-DEC-1999
		GB_RO:MMU6590	6429	AJ006590	AJ006590	6429	GB_RO:MMU6590	6429	AJ006590	Mus musculus mRNA for GANP protein.	38,678	2-Jun-99
		GB_HTG3:AC008852	116219	AC008852	AC008852	116219	GB_HTG3:AC008852	116219	AC008852	Homo sapiens chromosome 5 clone CITB-H1_2176121, *** SEQUENCING IN PROGRESS ***; 13 unordered pieces.	35,691	3-Aug-99
rx00169	2196	GB_GSS3:B11032	896	B11032	B11032	896	GB_GSS3:B11032	896	B11032	Arabidopsis thaliana genomic clone T17F10, genomic survey sequence.	42,024	14-MAY-1997
		GB_GSS3:B10120	909	B10120	B10120	909	GB_GSS3:B10120	909	B10120	T27N10-Sp6.1 TAMU Arabidopsis thaliana genomic clone T27N10, genomic survey sequence.	41,581	14-MAY-1997
		GB_GSS3:B09409	916	B09409	B09409	916	GB_GSS3:B09409	916	B09409	T27M2-Sp6 TAMU Arabidopsis thaliana genomic clone T27M2, genomic survey sequence.	41,356	14-MAY-1997
rx00170	1977	GB_GSS8:AQ027582	456	AQ027582	AQ027582	456	GB_GSS8:AQ027582	456	AQ027582	CIT-HSP-2325M20. TR CIT-HSP Homo sapiens genomic clone 2325M20, genomic survey sequence.	40,749	30-Jun-98
		GB_GSS6:AQ833529	484	AQ833529	AQ833529	484	GB_GSS6:AQ833529	484	AQ833529	HS_5304_B2_C02_TTA RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=880 Col=4 Row=F, genomic survey sequence.	38,017	27-Aug-99
rx00171	281	GB_PR2:HSN800174	2326	AL049389	AL049389	2326	GB_PR2:HSN800174	2326	AL049389	Homo sapiens mRNA; cDNA DKFZp586O0118 (from clone DKFZp586O0118).	37,556	21-MAY-1999
		GB_EST38:AL118463	279	AL118463	AL118463	279	GB_EST38:AL118463	279	AL118463	w9112a43 Beddington mouse dissected endoderm Mus musculus cDNA clone 528_12E22 5', mRNA sequence.	50,000	23-Sep-99
		GB_EST15:AA499834	419	AA499834	AA499834	419	GB_EST15:AA499834	419	AA499834	vg05e06.r1 Soares mouse NbMH Mus musculus cDNA clone IMAGE:860482 5', mRNA sequence.	39,801	1-Jul-97
rx00173	456	GB_EST24:AI211527	431	AI211527	AI211527	431	GB_EST24:AI211527	431	AI211527	p0h01a1.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone p0h01a1 5', mRNA sequence.	41,584	19-OCT-1998
		GB_PR3:AC004400	33367	AC004400	AC004400	33367	GB_PR3:AC004400	33367	AC004400	Homo sapiens chromosome 19, cosmid F24069, complete sequence.	38,902	12-MAR-1998

# **TABLE 4: ALIGNMENT RESULTS**

ductase precursor (fmr) mRNA, complete cds.  
 cosmid F24069, complete sequence.  
 glutinin-like protein 6 (ALS6) gene, complete cds.  
 RNA for chalcone synthase (CHS1).  
 RNA for chalcone synthase (CHS2).

oxidase mRNA, complete cds.  
 oxidase (AgProPO) gene, complete cds.

clone RP1-120G22, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in  
 clone RP1-120G22, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in  
 ondrion, complete genome.

drial genome.

ondrion, complete genome.

Homo sapiens cDNA clone IMAGE:2336756 3' similar to  
 DROXYINDOLE O-METHYLTRANSFERASE, mRNA sequence.  
 me survey sequence SP6 end of BAC BACN04K17 of DrosBAC  
 gaster (fruit fly), genomic survey sequence.  
 Homo sapiens cDNA clone IMAGE:2336756 3' similar to  
 DROXYINDOLE O-METHYLTRANSFERASE, mRNA sequence.  
 proved Human Genomic Sperm Library D Homo sapiens genomic  
 B, genomic survey sequence.  
 Homo sapiens genomic clone 2331N9, genomic survey sequence.

Homo sapiens cDNA clone IMAGE:2455893 3' similar to  
 DING PROTEIN, mRNA sequence.  
 ast BAC Library Magnaporthe grisea genomic clone mgxb0001M23f,  
 st BAC Library Magnaporthe grisea genomic clone mgxb0001P11f,

atrix metalloproteinase.  
 gitB and gitD genes for glutamine 2-oxoglutarate aminotransferase  
 plete cds.  
 mosome 2 clone BACR04107 (D644) RPCI-98 04.I.7 map 57B2-B3  
 NCING IN PROGRESS\*\*\*, 49 unordered pieces.  
 mosome 2 clone BACR04107 (D644) RPCI-98 04.I.7 map 57B2-B3  
 NCING IN PROGRESS\*\*\*, 49 unordered pieces.

Vicia faba	38,753	28-Sep-94
Homo sapiens	40,515	12-MAR-1998
Candida albicans	41,235	2-Jul-99
Juglans nigra x Juglans regia	39,558	19-Nov-99
Juglans nigra x Juglans regia	38,821	19-Nov-99
Corynebacterium glutamicum	100,000	8-Sep-97
Anopheles gambiae	35,751	23-Jul-98
Anopheles gambiae	36,395	5-Jan-99
Corynebacterium glutamicum	100,000	8-Sep-97
Homo sapiens	35,976	23-Nov-99
Homo sapiens	35,976	23-Nov-99
Mitochondrion Florometra serratissima	33,710	15-Jan-99
Mitochondrion Caenorhabditis elegans	35,036	30-Nov-97
Mitochondrion Florometra serratissima	36,021	15-Jan-99
Homo sapiens	37,800	2-Jun-99
Drosophila melanogaster	35,158	26-Jul-99
Homo sapiens	39,052	2-Jun-99
Homo sapiens	44,240	4-Nov-98
Homo sapiens	36,150	14-Jul-98
Homo sapiens	45,588	2-Sep-99
Magnaporthe grisea	63,529	06-MAR-1999
Magnaporthe grisea	49,580	06-MAR-1999
Bos taurus	40,440	30-MAR-1995
Corynebacterium glutamicum	36,237	13-MAR-1999
Drosophila melanogaster	36,045	2-Aug-99
Drosophila melanogaster	36,045	2-Aug-99

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

rx00200	3561	GB_BA2:MSU46844	16951	U46844	Mycobacterium smegmatis emB, embB, genes complete cds and putative propionyl-CoA carboxylase beta chain (pccB) genes, partial cds.	Mycobacterium smegmatis	53,937	12-MAY-1997
		GB_BA2:MAU66560	7853	U66560	Mycobacterium avium EmbR (embR), EmbA (embA) and EmbB (embB) genes, complete cds.	Mycobacterium avium	52,241	8-Nov-96
		GB_BA1:MTY13D12	37085	Z80343	Mycobacterium tuberculosis H37Rv complete genome; segment 156/162.	Mycobacterium tuberculosis	52,812	17-Jun-98
rx00207	441	GB_PR3:HTCRBCHR9	216293	AF029308	Homo sapiens chromosome 9 duplication of the T cell receptor beta locus and trypsinogen gene families.	Homo sapiens	39,286	13-Apr-98
		GB_PR3:HTCRBCHR9	216293	AF029308	Homo sapiens chromosome 9 duplication of the T cell receptor beta locus and trypsinogen gene families.	Homo sapiens	37,116	13-Apr-98
rx00211	786	GB_PR2:HSU81831	38674	U81831	Human cosmid LL12NCO1-67C6, ETv6 gene, intron 1A, partial sequence.	Homo sapiens	35,509	3-Jan-97
		GB_RO:MUSKROX252	2868	M28845	Mus musculus zinc finger protein (Krox-24) gene, exon 2.	Mus musculus	40,566	21-MAY-1996
		GB_HTG2:AC007440	120642	AC007440	Drosophila melanogaster chromosome 2 clone BACR37109 (D593) RPCI-98 37.1.9 map 49A-49B strain y; on bw sp, *** SEQUENCING IN PROGRESS ***; 103 unordered pieces.	Drosophila melanogaster	36,753	2-Aug-99
rx00218								
rx00220	627	GB_BA1:ASU04436	4668	U04436	Anabaena sp. PCC 7120 putative polyketide synthase gene, complete cds.	Anabaena sp.	33,766	21-DEC-1993
		GB_RO:AF068199	3490	AF068199	Mus musculus D-dopachrome tautomerase gene, complete cds.	Mus musculus	38,833	26-Aug-98
		GB_RO:AF068199	3490	AF068199	Mus musculus D-dopachrome tautomerase gene, complete cds.	Mus musculus	34,776	26-Aug-98
rx00222	1269	GB_PL1:AB011477	78181	AB011477	Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: MHK7, complete sequence.	Arabidopsis thaliana	36,766	20-Nov-99
		GB_EST17:AA615900	427	AA615900	vo91b05.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA clone IMAGE:1066449 5' similar to SW/MUCL_RAT P98089 INTESTINAL MUCIN-LIKE PROTEIN, mRNA sequence.	Mus musculus	39,782	07-OCT-1997
rx00230	843	GB_EST38:AW039188	486	AW039188	EST281423 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA clone cLET9F17, mRNA sequence.	Lycopersicon esculentum	41,286	18-OCT-1999
		GB_PR3:AC005255	94343	AC005255	Homo sapiens chromosome 19, CIT-HSP-146e8, complete sequence.	Homo sapiens	35,990	6-Jul-98
		GB_PR3:AC005255	94343	AC005255	Homo sapiens chromosome 19, CIT-HSP-146e8, complete sequence.	Homo sapiens	38,175	6-Jul-98
rx00232	633	GB_HTG2:AC007118	200000	AC007118	Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***; 45 unordered pieces.	Homo sapiens	36,772	19-MAR-1999
		GB_HTG2:AC007118	200000	AC007118	Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***; 45 unordered pieces.	Homo sapiens	36,772	19-MAR-1999
		GB_GSS1:CNS004WZ	978	AL055537	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR11G02 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	Drosophila melanogaster	34,518	3-Jun-99
rx00233	517	GB_BA1:AB006206	7443	AB006206	Streptomyces griseus AmfR, AmfA and AmfB genes and 4 ORFs, complete cds.	Streptomyces griseus	38,690	5-Feb-99
		GB_PR4:AC006999	112878	AC006999	Homo sapiens clone NH0462A19, complete sequence.	Homo sapiens	40,244	17-Jul-99
		GB_HTG2:AC007042	132400	AC007042	Homo sapiens clone NH0399H17, *** SEQUENCING IN PROGRESS ***; 5 unordered pieces.	Homo sapiens	40,244	06-MAR-1999
rx00234	653	GB_PAT:E13059	3480	E13059	gDNA encoding cytochrome b5.	unidentified	40,091	24-Jun-98
		GB_PL1:AB022444	2104	AB022444	Monterella alpina gene for cytochrome b5, complete cds.	Monterella alpina	42,314	14-Jul-99
		GB_GSS9:AAQ112619	443	AAQ112619	CIT-HSP-2371D11, TR CIT-HSP Homo sapiens genomic clone 2371D11, genomic survey sequence.	Homo sapiens	39,623	29-Aug-98
rx00236	849	GB_BA1:CGPROMF34	60	X90361	C.glutamicum DNA for promoter fragment F34.	Corynebacterium glutamicum	98,333	4-Nov-96
		GB_IN1:CEF56G4	38062	Z81552	Caenorhabditis elegans cosmid F56G4, complete sequence.	Caenorhabditis elegans	36,190	08-OCT-1999
		GB_EST16:C51159	370	C51159	C51159 Yujii Kohara unpublished cDNA: Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk49h3 5', mRNA sequence.	Caenorhabditis elegans	41,096	18-OCT-1999
rx00237	501	GB_GSS9:AAQ148605	511	AQ148605	HS_3137_B2_A11_17 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-3137 Col=22 Row=B, genomic survey sequence.	Homo sapiens	37,959	08-OCT-1998
		GB_GSS11:AAQ274889	622	AQ274889	RPCI-5-1111N8T7 RPCI-5 Homo sapiens genomic clone RPCI-5-1111N8T7, genomic survey sequence.	Homo sapiens	44,628	10-Nov-98

**TABLE 4: ALIGNMENT RESULTS**

rx00238	492	GB_GSS11:AQ274889	622	AQ274889	RPCL5-1111N8T7 RPCI-5 Homo sapiens genomic clone RPCI-5-1111N8T7, genomic survey sequence.	Homo sapiens	37,321	10-Nov-98
		GB_PL2:PB091560	2605	U91560	Paracoccidioides brasiliensis heat shock protein 70 (Hsp70) gene, complete cds.	Paracoccidioides brasiliensis	37,137	12-MAR-1999
		GB_HTG6:AC007857	212658	AC007857	Homo sapiens, *** SEQUENCING IN PROGRESS ***; 2 ordered pieces.	Homo sapiens	38,285	26-Nov-99
		GB_PR4:AC009288	140876	AC009288	Homo sapiens, complete sequence.	Homo sapiens	36,575	19-Nov-99
rx00239	708	GB_PL2:PB091560	2605	U91560	Paracoccidioides brasiliensis heat shock protein 70 (Hsp70) gene, complete cds.	Paracoccidioides brasiliensis	45,545	12-MAR-1999
		GB_BA2:CJ096452	1450	U96452	Campylobacter jejuni major outer membrane porin gene, complete cds.	Campylobacter jejuni	37,197	02-DEC-1998
		GB_RO:RATPF4	1675	M15254	Rat platelet factor 4 (PF4) gene.	Rattus norvegicus	35,014	27-Apr-93
rx00240	333	GB_BA1:CGLYSI	4232	X60312	C.glutamicum lysI gene for L-lysine permease.	Corynebacterium glutamicum	46,207	30-Jan-92
		GB_PR3:AC005358	184886	AC005358	Homo sapiens chromosome 17, clone hRPK.746_E_8, complete sequence.	Homo sapiens	35,843	29-Aug-98
		GB_BA1:CGLYSI	4232	X60312	C.glutamicum lysI gene for L-lysine permease.	Corynebacterium glutamicum	42,025	30-Jan-92
rx00242	1401	GB_BA1:CGLYSI	4232	X60312	C.glutamicum lysI gene for L-lysine permease.	Corynebacterium glutamicum	100,000	30-Jan-92
		GB_PR3:HSJ514B11	100494	AL049554	Human DNA sequence from clone 514B11 on chromosome 6q16.1-21 Contains an EST, STS and GSSs, complete sequence.	Homo sapiens	37,010	23-Nov-99
		GB_HTG3:AC009393	137353	AC009393	Drosophila melanogaster chromosome 3 clone BACR17F05 (D977) RPCI-98 17.F.5 map 87D-87D strain y, cn bw sp, *** SEQUENCING IN PROGRESS ***; 111 unordered pieces.	Drosophila melanogaster	39,600	27-Aug-99
rx00244	759	GB_HTG4:AC011290	148409	AC011290	Homo sapiens clone NH006402, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Homo sapiens	38,102	15-OCT-1999
		GB_HTG4:AC011290	148409	AC011290	Homo sapiens clone NH006402, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Homo sapiens	38,102	15-OCT-1999
		GB_EST23:A077162	527	A077162	TENU3384 T. cruzi epimastigote normalized cDNA Library Trypanosoma cruzi cDNA clone 2806 5' similar to TRANSPLANTATION ANTIGEN P35B spIP23591TUM3_MOUSE, mRNA sequence.	Trypanosoma cruzi	38,847	10-Aug-98
rx00245	1608	GB_HTG2:AC007644	141048	AC007644	Homo sapiens chromosome 17 clone hRPK.19_F_16 map 17, *** SEQUENCING IN PROGRESS ***; 17 unordered pieces.	Homo sapiens	36,929	23-MAY-1999
		GB_HTG2:AC007644	141048	AC007644	Homo sapiens chromosome 17 clone hRPK.19_F_16 map 17, *** SEQUENCING IN PROGRESS ***; 17 unordered pieces.	Homo sapiens	36,929	23-MAY-1999
		GB_HTG2:AC007644	141048	AC007644	Homo sapiens chromosome 17 clone hRPK.19_F_16 map 17, *** SEQUENCING IN PROGRESS ***; 17 unordered pieces.	Homo sapiens	34,025	23-MAY-1999
rx00247	1050	GB_PR2:AP000119	100000	AP000119	Homo sapiens genomic DNA of 21q22.1, GART and AML related, SLC5A3-4A4 region, segment 2/8, complete sequence.	Homo sapiens	36,187	25-Sep-99
		GB_PR2:AP000051	100000	AP000051	Homo sapiens genomic DNA, chromosome 21q22.1, segment 22/28, complete sequence.	Homo sapiens	36,187	20-Nov-99
		GB_PR2:AP000166	100000	AP000166	Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone B2344F14-150E8, segment 2/9, complete sequence.	Homo sapiens	37,942	20-Nov-99
rx00248	846	GB_PR4:AC006464	99808	AC006464	Homo sapiens BAC clone NH0436C12 from 2, complete sequence.	Homo sapiens	36,797	22-OCT-1999
		GB_PR4:AC006238	211945	AC006238	Homo sapiens chromosome 18, clone hRPK.474_N_24, complete sequence.	Homo sapiens	32,896	31-Jan-99
		GB_PR4:AC006238	211945	AC006238	Homo sapiens chromosome 18, clone hRPK.474_N_24, complete sequence.	Homo sapiens	34,438	31-Jan-99
rx00250	870	GB_GSS10:AQ244736	469	AQ244736	HS_2056_B1_F03_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2056 Col=5 Row=L, genomic survey sequence.	Homo sapiens	36,310	03-OCT-1998
		GB_PAT:107323	340	I07323	Sequence 5 from Patent EP 0331961.	Unknown	38,125	02-DEC-1994
		GB_PR1:HS11Q13RP	10777	Y12377	H.sapiens FGF/int-2 gene upstream flanking region.	Homo sapiens	36,155	17-Apr-97
rx00252	366	GB_BA1:MTCY20G9	37218	Z77162	Mycobacterium tuberculosis H37Rv complete genome; segment 25/162.	Mycobacterium tuberculosis	39,554	17-Jun-98
		GB_BA1:MTV004	69350	AL009198	Mycobacterium tuberculosis H37Rv complete genome; segment 144/162.	Mycobacterium tuberculosis	40,443	18-Jun-98
		GB_BA1:MTV004	69350	AL009198	Mycobacterium tuberculosis H37Rv complete genome; segment 144/162.	Mycobacterium tuberculosis	41,803	18-Jun-98

**TABLE 4: ALIGNMENT RESULTS**

rx00256	894	GB_PR4:AC005343	137213	AC005343	Homo sapiens Chromosome 12p13.3 BAC RPC11-21K20 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	Homo sapiens	36,436	2-Apr-99
		GB_PR3:AC003005	45084	AC003005	Human DNA from chromosome 19-specific cosmid F25419 containing ZNF gene family members, genomic sequence, complete sequence.	Homo sapiens	36,395	22-OCT-1997
		GB_HTG3:AC007930	67668	AC007930	Drosophila melanogaster chromosome 2 clone BACR49A06 (D772) RPC1-98 49 A.6 map 43B-43B strain Y; on bw sp, *** SEQUENCING IN PROGRESS***, 40 unordered pieces.	Drosophila melanogaster	32,503	20-Sep-99
rx00257	579	GB_PR4:AC005343	137213	AC005343	Homo sapiens Chromosome 12p13.3 BAC RPC11-21K20 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	Homo sapiens	37,063	2-Apr-99
		GB_HTG1:HS1096J16	194423	AL121721	Homo sapiens chromosome 20 clone RP5-1096J16, *** SEQUENCING IN PROGRESS***, in unordered pieces.	Homo sapiens	37,217	23-Nov-99
		GB_HTG1:HS1096J16	194423	AL121721	Homo sapiens chromosome 20 clone RP5-1096J16, *** SEQUENCING IN PROGRESS***, in unordered pieces.	Homo sapiens	37,217	23-Nov-99
rx00258	795	GB_PR3:HSJ747H23	114201	AL049899	Human DNA sequence from clone 747H23 on chromosome 6q13-15. Contains the 3' part of the ME1 gene for malic enzyme 1, soluble (NADP-dependent malic enzyme, malate oxidoreductase, EC 1.1.1.40), a novel gene and the 5' part of the gene for N-acetylglucosamine-phosphate mutase. Contains ESTs, STSS, GSSs and two putative CpG islands, complete sequence.	Homo sapiens	36,469	23-Nov-99
		GB_HTG2:HSJ202D23	175496	AL121716	Homo sapiens chromosome 6 clone RP1-202D23 map q14.1-15, *** SEQUENCING IN PROGRESS***, in unordered pieces.	Homo sapiens	36,469	03-DEC-1999
		GB_HTG2:HSJ202D23	175496	AL121716	Homo sapiens chromosome 6 clone RP1-202D23 map q14.1-15, *** SEQUENCING IN PROGRESS***, in unordered pieces.	Homo sapiens	36,469	03-DEC-1999
rx00260	1299	GB_PR3:HS360B4	23388	AL031716	Human DNA sequence from clone 360B4 on chromosome 16. Contains part of a gene for a PUTATIVE novel protein similar to predicted bacterial and worm proteins and ESTs, complete sequence.	Homo sapiens	36,145	23-Nov-99
		GB_EST19:AA741904	423	AA741904	LmLV39p371A Leishmania major promastigote full length cDNA library from early logarithmic stage (day 3) Leishmania major cDNA clone 71A 5' mRNA sequence.	Leishmania major	39,192	10-DEC-1998
		GB_PR3:HS360B4	23388	AL031716	Human DNA sequence from clone 360B4 on chromosome 16. Contains part of a gene for a PUTATIVE novel protein similar to predicted bacterial and worm proteins and ESTs, complete sequence.	Homo sapiens	39,494	23-Nov-99
rx00264								
rx00267	441	GB_GSS10:AQ258013	761	AQ258013	nbxb0019H05f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0019H05f, genomic survey sequence.	Oryza sativa	52,033	23-OCT-1998
		GB_EST10:AA167894	552	AA167894	CpEST_021 unizAPCpQWASporLib1 Cryptosporidium parvum cDNA 5' similar to lactate dehydrogenase, mRNA sequence.	Cryptosporidium parvum	40,462	19-DEC-1996
		GB_HTG3:AC011591	129431	AC011591	Homo sapiens chromosome 17 clone 118_B_18 map 17, *** SEQUENCING IN PROGRESS***, 25 unordered pieces.	Homo sapiens	35,469	07-OCT-1999
rx00271	1113	GB_PL1:CHTRP1	3480	X70035	C.heterostrophus gene for trifunctional tryptophan synthase.	Cochliobolus heterostrophus	41,636	31-OCT-1996
		GB_Vi:FLU47643	1492	U47643	Feline leukemia virus Notch2 gene, clone FelV/Notch2-AP (subgenomic), partial cds.	Feline leukemia virus	37,869	25-OCT-1996
		GB_Vi:FLU47644	1641	U47644	Feline leukemia virus Notch2 gene, clone FelV/Notch2-B, partial cds.	Feline leukemia virus	36,441	25-OCT-1996
rx00272	495	GB_GSS8:AQ041841	373	AQ041841	CIT-HSP-2335L1 TR CIT-HSP Homo sapiens genomic clone 2335L1, genomic survey sequence.	Homo sapiens	45,455	14-Jul-98
		GB_GSS13:AQ429301	591	AQ429301	CITBI-E1-2562H16 TR CITBI-E1 Homo sapiens genomic clone 2562H16, genomic survey sequence.	Homo sapiens	63,636	24-MAR-1999
		GB_GSS10:AQ237541	667	AQ237541	RPC111-61O21.TJB.1 RPC1-11 Homo sapiens genomic clone RPC1-11-61O21, genomic survey sequence.	Homo sapiens	62,222	21-Apr-99
rx00273	1236	GB_BA1:CGBETPGEN	2339	X93514	C-glutamicum bclP gene.	Corynebacterium glutamicum	44,056	8-Sep-97

TABLE 4: ALIGNMENT RESULTS

GB_PR2:HS142F18	141672	AL031073	Human DNA sequence from clone 142F18 on chromosome Xq26.3-27.2 Contains part of a gene similar to melanoma-associated antigen, EST, GSS and an inverted repeat, complete sequence.	Homo sapiens	44,543	23-Nov-99
GB_IN2:AC007177	101320	AC007177	Drosophila melanogaster, chromosome 2R, region 59C1-59C5, P1 clones DS06621 and DS02186, complete sequence.	Drosophila melanogaster	36,721	27-MAR-1999
GB_HTG3:AC011675	98026	AC011675	Homo sapiens clone 10_J_17, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	35,405	10-OCT-1999
GB_HTG3:AC011675	98026	AC011675	Homo sapiens clone 10_J_17, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	35,405	10-OCT-1999
GB_HTG3:AC010598	174019	AC010598	Homo sapiens chromosome 5 clone CIT-HSPC_5609, *** SEQUENCING IN PROGRESS *** , 50 unordered pieces.	Homo sapiens	36,908	16-Sep-99
GB_GSS14:AQ574926	666	AQ574926	nbxb0086K14f CUGI Rice BAC Library Onyza sativa genomic clone nbxb0086K14f, genomic survey sequence.	Onyza sativa	33,830	2-Jun-99
GB_HTG2:AC004396	43686	AC004396	Homo sapiens, *** SEQUENCING IN PROGRESS *** , 2 unordered pieces.	Homo sapiens	38,298	19-Jul-99
GB_HTG2:AC004396	43686	AC004396	Homo sapiens, *** SEQUENCING IN PROGRESS *** , 2 unordered pieces.	Homo sapiens	38,298	19-Jul-99
GB_PL1:SC9952X	29286	Z49212	S cerevisiae chromosome XIII cosmid 9952.	Saccharomyces cerevisiae	37,118	11-Aug-97
GB_PL1:S45357	4017	S45357	PSE-1=protein secretion enhancer [Saccharomyces cerevisiae, Genomic, 4017 nt].	Saccharomyces cerevisiae	41,394	08-MAY-1993
GB_PL1:SCPSE1G	4017	Z11538	S. cerevisiae PSE-1 gene.	Saccharomyces cerevisiae	41,394	13-Aug-96
GB_HTG3:AC009911	99707	AC009911	Drosophila melanogaster, chromosome 2 clone BACR01N17 (D1036) RPCI-98 01.N.17 map 38A-38A strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** , 69 unordered pieces.	Drosophila melanogaster	33,023	05-OCT-1999
GB_HTG3:AC009911	99707	AC009911	Drosophila melanogaster, chromosome 2 clone BACR01N17 (D1036) RPCI-98 01.N.17 map 38A-38A strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** , 69 unordered pieces.	Drosophila melanogaster	33,023	05-OCT-1999
GB_HTG4:AC008397	230451	AC008397	Homo sapiens chromosome 19 clone CIT-HSPC_251H24, *** SEQUENCING IN PROGRESS *** , 81 unordered pieces.	Homo sapiens	37,432	31-OCT-1999
GB_PR3:HSB11B7	37290	Z82171	Human DNA sequence from cosmid B11B7 on chromosome 22 contains ESTs.	Homo sapiens	35,845	23-Nov-99
GB_PR3:HSB11B7	37290	Z82171	Human DNA sequence from cosmid B11B7 on chromosome 22 contains ESTs.	Homo sapiens	37,126	23-Nov-99
GB_RO:RATMTA	4197	L39264	Rattus norvegicus beta-2 adrenergic receptor gene, complete cds and promoter region.	Rattus norvegicus	41,259	23-Feb-96
GB_GSS8:AQ030327	411	AQ030327	HS_2177_B1_H06 MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2177 Col=11 Row=P, genomic survey sequence.	Homo sapiens	37,656	1-Jul-98
GB_PR3:HSL118GB	27858	Z68883	Human DNA sequence from cosmid L118G10, Huntington's Disease Region, chromosome 4p16.3.	Homo sapiens	37,412	23-Nov-99
GB_PR3:HSJ513G18	110770	AL109760	Human DNA sequence from clone 513G18 on chromosome 4, complete sequence.	Homo sapiens	37,412	23-Nov-99
GB_EST10:AA157040	414	AA157040	zo51c05.r1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone IMAGE:590408 5' similar to gb:M84711 40S RIBOSOMAL PROTEIN S3A (HUMAN);, mRNA sequence.	Homo sapiens	37,136	11-DEC-1996
GB_EST11:AA213935	629	AA213935	zn57a04.r1 Stratagene muscle 937209 Homo sapiens cDNA clone IMAGE:562254 5' similar to gb:M84711 40S RIBOSOMAL PROTEIN S3A (HUMAN);, mRNA sequence.	Homo sapiens	34,219	1-Aug-97
GB_STS:BLYBG	459	L43987	Hordeum vulgare (clone ABG380) chromosome 4H.6H.7H STS mRNA, sequence tagged site.	Hordeum vulgare	37,786	27-Jul-95
GB_PR2:HSAC000121	93163	AC000121	Human BAC clone RC249A12 from Tq22, complete sequence.	Homo sapiens	36,735	31-Jan-97
GB_BA2:CGU31281	1614	U31281	Corynebacterium glutamicum biotin synthase (bioB) gene, complete cds.	Corynebacterium glutamicum	100,000	21-Nov-96
GB_PR2:HSAC000121	93163	AC000121	Human BAC clone RC249A12 from Tq22, complete sequence.	Homo sapiens	37,662	31-Jan-97
GB_HTG2:AC006938	82665	AC006938	Drosophila melanogaster chromosome 2 clone DS01630 (D506) map 60C7-60C8 strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** , 9 unordered pieces.	Drosophila melanogaster	37,241	2-Aug-99
GB_HTG2:AC007116	25478	AC007116	Drosophila melanogaster chromosome 2 clone DS04467 (D447) map 60C6-60C8 strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** , 5 unordered pieces.	Drosophila melanogaster	38,630	30-Jul-99
GB_HTG2:AC006938	82665	AC006938	Drosophila melanogaster chromosome 2 clone DS01630 (D506) map 60C7-60C8 strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** , 9 unordered pieces.	Drosophila melanogaster	37,241	2-Aug-99
GB_GSS14:AQ585202	564	AQ585202	RPCI-11-451L11, TJ RPCI-11 Homo sapiens genomic clone RPCI-11-451L11, genomic survey sequence.	Homo sapiens	37,319	7-Jun-99
GB_BA1:NGPILC1	3144	Y13022	N gonorrhoeae pilC1 gene, strain 640.	Neisseria gonorrhoeae	38,667	07-OCT-1997



TABLE 4: ALIGNMENT RESULTS

	GB_BA1:NGPILC1	3144	Y13022	N.gonorrhoeae pilC1 gene, strain 640.	Neisseria gonorrhoeae	36,000	07-OCT-1997
rx00321							
rx00322	1227	153053	AC007533	Homo sapiens chromosome 16 clone 474B12, *** SEQUENCING IN PROGRESS ***; 5 ordered pieces.	Homo sapiens	39,469	12-MAY-1999
		153053	AC007533	Homo sapiens chromosome 16 clone 474B12, *** SEQUENCING IN PROGRESS ***; 5 ordered pieces.	Homo sapiens	39,469	12-MAY-1999
rx00325	768	82512	AC000061	Human BAC clone 133K23 from 7q31.2, complete sequence.	Homo sapiens	38,950	14-Nov-96
		1273	U73860	Corynebacterium diphtheriae heme oxygenase homolog (hmuO) gene, complete cds.	Corynebacterium diphtheriae	52,604	7-Feb-97
		652	AB019621	Corynebacterium diphtheriae mRNA for Heme oxygenase, complete cds.	Corynebacterium diphtheriae	55,675	31-Jul-99
		554	AI096171	28 EcoRI Rice Etiolated Leaf cDNA Library Oriza sativa cDNA clone RZ513, mRNA sequence.	Oriza sativa	38,536	19-Aug-98
rx00326	603	11624	X51522	Bacteriophage P4 complete DNA genome.	Bacteriophage P4	40,577	17-Feb-97
		3063	X05623	Bacteriophage P4 alpha gene and cis replication region crr.	Bacteriophage P4	38,640	12-Sep-93
		30480	Z68336	Caenorhabditis elegans cosmid F22B3, complete sequence.	Caenorhabditis elegans	39,012	2-Sep-99
rx00334	459	3686	Y13221	Corynebacterium glutamicum glnA gene.	Corynebacterium glutamicum	37,640	28-Aug-97
		259	AQ248516	T5J22-Sp6 TAMU Arabidopsis thaliana genomic clone T5J22, genomic survey sequence.	Arabidopsis thaliana	38,525	06-OCT-1998
		3686	Y13221	Corynebacterium glutamicum glnA gene.	Corynebacterium glutamicum	40,487	28-Aug-97
rx00336	594	3686	Y13221	Corynebacterium glutamicum glnA gene.	Corynebacterium glutamicum	34,797	28-Aug-97
		7026	X95549	X.laevis FIM-B.1 gene.	Xenopus laevis	33,217	13-Feb-97
		3686	Y13221	Corynebacterium glutamicum glnA gene.	Corynebacterium glutamicum	37,371	28-Aug-97
rx00337	1173	3464	U43536	Corynebacterium glutamicum heat shock, ATP-binding protein (cipB) gene, complete cds.	Corynebacterium glutamicum	36,406	13-MAR-1997
		1160	AJ004934	Corynebacterium glutamicum dapD gene, complete CDS.	Corynebacterium glutamicum	39,734	17-Jun-98
		396	AI944838	bs06a08.y1 Drosophila melanogaster adult testis library Drosophila melanogaster cDNA clone bs06a08.5', mRNA sequence.	Drosophila melanogaster	37,626	17-Aug-99
rx00338	1263	7725	X04960	Brevibacterium lactofermentum tryptophan operon.	Corynebacterium glutamicum	39,790	10-Feb-99
		7725	E01688	Genomic DNA of trp operon of prebacterium latophelmentam.	unidentified	39,871	29-Sep-97
		7726	E01375	DNA sequence of tryptophan operon.	Corynebacterium glutamicum	39,871	29-Sep-97
rx00339	840	131993	U75930	Orygia pseudotsugata nuclear polyhedrosis virus complete genome.	Orygia pseudotsugata nuclear polyhedrosis virus	38,264	06-MAR-1998
		131993	U75930	Orygia pseudotsugata nuclear polyhedrosis virus complete genome.	Orygia pseudotsugata nuclear polyhedrosis virus	38,620	06-MAR-1998
rx00342	552	126593	AC008340	Drosophila melanogaster chromosome 2 clone BACR07J20 (D918) RPCI-98 07.J.20 map 42D-42E strain y, cn bw sp *** SEQUENCING IN PROGRESS ***; 92 unordered pieces.	Drosophila melanogaster	38,193	6-Aug-99
		636	AI794353	fc43c12.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA 5' similar to TR:Q62868 Q62868 ROK-ALPHA.1, mRNA sequence.	Danio rerio	40,283	2-Jul-99
		33676	U73635	Human Chromosome 11 Cosmid cSRL156b6, complete sequence.	Homo sapiens	39,366	25-Jul-97
		33676	U73635	Human Chromosome 11 Cosmid cSRL156b6, complete sequence.	Homo sapiens	37,970	25-Jul-97
rx00344	1002	33595	AL032819	Homo sapiens chromosome 16 clone LA16-312E8. *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	40,415	03-DEC-1999

TABLE 4: ALIGNMENT RESULTS

GB_HTG2:HS312E8	33595	AL032819	Homo sapiens chromosome 16 clone LA16-312E8, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	40,415	03-DEC-1999
GB_OM:BOVINOPHOS	1573	M55916	Bovine inositol polyphosphate 1-phosphatase (inositol polyphosphate 1-phosphatase gene) mRNA, complete cds.	Bos taurus	38,420	27-Apr-93
GB_PR3:HS353E16	189765	AL031591	Human DNA sequence from clone 353E16 on chromosome 22q11.22-12.3, complete sequence.	Homo sapiens	34,766	23-Nov-99
GB_HTG2:AC005059	170128	AC005059	Homo sapiens clone RG07AA24, *** SEQUENCING IN PROGRESS ***; 25 unordered pieces.	Homo sapiens	37,011	13-MAR-1999
GB_HTG2:AC005059	170128	AC005059	Homo sapiens clone RG07AA24, *** SEQUENCING IN PROGRESS ***; 25 unordered pieces.	Homo sapiens	37,011	13-MAR-1999
GB_BA1:D87976	2352	D87976	Brevibacterium lactofermentum DNA for D-2-hydroxyisocaproate dehydrogenase (ddh), complete cds.	Corynebacterium glutamicum	39,290	7-Feb-99
GB_BA1:CGDDH	1829	Y00151	Corynebacterium glutamicum ddh gene for meso-diaminopimelate D-dehydrogenase (EC 1.4.1.16).	Corynebacterium glutamicum	39,342	12-Sep-93
GB_BA1:CGDDH	1829	Y00151	Corynebacterium glutamicum ddh gene for meso-diaminopimelate D-dehydrogenase (EC 1.4.1.16).	Corynebacterium glutamicum	38,624	12-Sep-93
GB_BA1:CGDDH	1829	Y00151	Corynebacterium glutamicum ddh gene for meso-diaminopimelate D-dehydrogenase (EC 1.4.1.16).	Corynebacterium glutamicum	100,000	12-Sep-93
GB_BA1:D87976	2352	D87976	Brevibacterium lactofermentum DNA for D-2-hydroxyisocaproate dehydrogenase (ddh), complete cds.	Corynebacterium glutamicum	98,411	7-Feb-99
GB_PAT:E14511	1034	E14511	DNA encoding Brevibacterium diaminopimelic acid dehydrogenase.	Corynebacterium glutamicum	100,000	28-Jul-99
GB_HTG4:AC009043	170748	AC009043	Homo sapiens chromosome 16 clone RPC1-11_184F14, *** SEQUENCING IN PROGRESS ***; 122 Homo sapiens unordered pieces.	122 Homo sapiens	37,337	31-OCT-1999
GB_HTG4:AC009043	170748	AC009043	Homo sapiens chromosome 16 clone RPC1-11_184F14, *** SEQUENCING IN PROGRESS ***; 122 Homo sapiens unordered pieces.	122 Homo sapiens	37,337	31-OCT-1999
GB_PR4:HSZ02JP09	811	AF177521	Homo sapiens tight junction protein ZO-2 (TJP2) gene, exons 8 and 9.	Homo sapiens	40,758	28-Sep-99
GB_PAT:AR004983	2277	AR004983	Sequence 5 from patent US 5747317.	Unknown.	41,638	04-DEC-1998
GB_EST37:AI967505	380	AI967505	Ljimpes03-215-c10 Ljimp Lambda HybriZap two-hybrid library Lotus japonicus cDNA clone LP215-03-c10 5' similar to 60S ribosomal protein L39, mRNA sequence.	Lotus japonicus	45,882	24-Aug-99
GB_EST27:AI399460	670	AI399460	NCSF4F6T7 Subtracted Perithecial Neurospora crassa cDNA clone SP4F6 3', mRNA sequence.	Neurospora crassa	38,571	8-Feb-99
GB_IN2:AC004445	61852	AC004445	Drosophila melanogaster DNA sequence (P1 DS00445 (D93)), complete sequence.	Drosophila melanogaster	37,236	01-MAY-1998
GB_HTG6:AC011694	160557	AC011694	Homo sapiens clone RP11-19D19, *** SEQUENCING IN PROGRESS ***; 33 unordered pieces.	Homo sapiens	34,087	03-DEC-1999
GB_HTG6:AC011694	160557	AC011694	Homo sapiens clone RP11-19D19, *** SEQUENCING IN PROGRESS ***; 33 unordered pieces.	Homo sapiens	40,523	03-DEC-1999
GB_BA1:COXHSPAB	2302	M20482	C.burnetii heat shock operon encoding two heat shock proteins (hspA and hspB), complete cds.	Coxiella burnetii	37,788	26-Apr-93
GB_RO:CBGPIMR	1735	Z37977	C.barabensis (griseus) mRNA for glucose phosphate isomerase.	Cricetulus griseus	37,823	14-Sep-95
GB_GSS10:AQ172617	505	AQ172617	HS_3197_A2_G09_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3197 Col=18 Row=M, genomic survey sequence.	Homo sapiens	37,580	17-OCT-1998
GB_BA1:MTY25D10	40838	Z95558	Mycobacterium tuberculosis H37Rv complete genome; segment 28/162.	Mycobacterium tuberculosis	60,477	17-Jun-98
GB_BA1:MSGY224	40051	AD0000004	Mycobacterium tuberculosis sequence from clone y224.	Mycobacterium tuberculosis	60,270	03-DEC-1996
GB_BA1:U00018	42991	U00018	Mycobacterium leprae cosmid B2168.	Mycobacterium leprae	37,913	01-MAR-1994
GB_IN1:DMU29153	8230	U29153	Drosophila melanogaster nudel (ndt) mRNA, complete cds.	Drosophila melanogaster	36,190	08-DEC-1995
GB_IN1:DMU29153	8230	U29153	Drosophila melanogaster nudel (ndt) mRNA, complete cds.	Drosophila melanogaster	37,202	08-DEC-1995
GB_IN2:AE001274	268984	AE001274	Leishmania major chromosome 1, complete sequence.	Leishmania major	37,885	24-MAR-1999
GB_EST11:AA270543	516	AA270543	va88h06.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone IMAGE:736571 5', mRNA sequence.	Mus musculus	40,546	26-MAR-1997
GB_IN2:AE001274	268984	AE001274	Leishmania major chromosome 1, complete sequence.	Leishmania major	36,103	24-MAR-1999
GB_GSS9:AQ158990	728	AQ158990	nbxb0012L11r CUGI Rice BAC Library Oryza sativa genomic clone nbxb0012L11r, genomic survey sequence.	Oryza sativa	41,463	12-Sep-98

TABLE 4: ALIGNMENT RESULTS

GB_GSS12:AQ342952	761	AQ342952	RPC111-122O15.TV RPC11-11 Homo sapiens genomic clone RPC11-1-122O15, genomic survey sequence.	Homo sapiens	37,556	07-MAY-1999
GB_GSS9:AQ158990	728	AQ158990	nbxb0012L11r CUGI Rice BAC Library Oryza sativa genomic clone nbxb0012L11r, genomic survey sequence.	Oryza sativa	37,923	12-Sep-98
GB_PR2:HS1052M9	134245	AL022718	Human DNA sequence from clone 1052M9 on chromosome Xq25. Contains the SH2D1A gene for SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) (DSHP), part of a 60S Acidic Ribosomal protein 1 (RPLP1) LIKE gene and part of a mouse DOC4 LIKE gene. Contains ESTs and GSSs, complete sequence.	Homo sapiens	43,564	23-Nov-99
GB_BA2:RCPHSYNG	45959	Z11165	R capsulatus complete photosynthesis gene cluster.	Rhodobacter capsulatus	36,930	2-Sep-99
GB_PR2:HS1052M9	134245	AL022718	Human DNA sequence from clone 1052M9 on chromosome Xq25. Contains the SH2D1A gene for SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) (DSHP), part of a 60S Acidic Ribosomal protein 1 (RPLP1) LIKE gene and part of a mouse DOC4 LIKE gene. Contains ESTs and GSSs, complete sequence.	Homo sapiens	33,981	23-Nov-99
GB_PL2:AC002311	85855	AC002311	Arabidopsis thaliana chromosome I BAC T26J12 genomic sequence, complete sequence.	Arabidopsis thaliana	38,957	4-Feb-98
GB_PL2:AC002311	85855	AC002311	Arabidopsis thaliana chromosome I BAC T26J12 genomic sequence, complete sequence.	Arabidopsis thaliana	36,300	4-Feb-98
GB_PL1:HASMT	27694	D31785	Hansenula wingei mitochondrial DNA, complete sequence.	Mitochondrion Pichia canadensis	33,985	10-Jun-99
GB_PL1:HASMT	27694	D31785	Hansenula wingei mitochondrial DNA, complete sequence.	Mitochondrion Pichia canadensis	33,775	10-Jun-99
GB_HTG1:CNS01DRT	222150	AL118557	Homo sapiens chromosome 14 clone R-1033H12, *** SEQUENCING IN PROGRESS ***, in ordered pieces.	Homo sapiens	35,417	25-Sep-99
GB_HTG1:CNS01DRT	222150	AL118557	Homo sapiens chromosome 14 clone R-1033H12, *** SEQUENCING IN PROGRESS ***, in ordered pieces.	Homo sapiens	35,417	25-Sep-99
GB_HTG1:CNS01DRT	222150	AL118557	Homo sapiens chromosome 14 clone R-1033H12, *** SEQUENCING IN PROGRESS ***, in ordered pieces.	Homo sapiens	38,005	25-Sep-99
GB_IN1:CELC32B5	42545	U80843	Caenorhabditis elegans cosmid C32B5.	Caenorhabditis elegans	36,468	05-DEC-1996
GB_IN1:CELC32B5	42545	U80843	Caenorhabditis elegans cosmid C32B5.	Caenorhabditis elegans	40,000	05-DEC-1996
GB_PR4:AC005940	158414	AC005940	Homo sapiens chromosome 17, clone HRPK.167_N_20, complete sequence.	Homo sapiens	35,219	18-MAR-1999
GB_PR1:HSSERCA11	3050	Y15726	Homo sapiens SERCA3 gene, exons 11-14.	Homo sapiens	35,036	30-Jun-98
GB_PR4:AC005940	158414	AC005940	Homo sapiens chromosome 17, clone HRPK.167_N_20, complete sequence.	Homo sapiens	36,826	18-MAR-1999
GB_PR1:HSSERCA11	3050	Y15726	Homo sapiens SERCA3 gene, exons 11-14.	Homo sapiens	38,555	30-Jun-98
GB_PR4:AC005940	158414	AC005940	Homo sapiens chromosome 17, clone HRPK.167_N_20, complete sequence.	Homo sapiens	36,370	18-MAR-1999
GB_PR3:HS591B8	142552	AL035410	Human DNA sequence from clone 591B8 on chromosome 1p13.1, complete sequence.	Homo sapiens	35,891	23-Nov-99
GB_BA1:AP000003	233000	AP000003	Pyrococcus horikoshii OT3 genomic DNA, 544001-777000 nt, position (3/7).	Pyrococcus horikoshii	36,849	8-Feb-99
GB_PL2:F25A4	115721	AC008263	Arabidopsis thaliana chromosome 1 BAC F25A4 sequence, complete sequence.	Arabidopsis thaliana	37,628	15-Sep-99
GB_PR3:HS413H6	142599	AL022724	Human DNA sequence from clone 413H6 on chromosome 6p22.3-24.3. Contains a hamster Androgen-dependent Expressed Protein like protein gene. ESTs and GSSs, complete sequence.	Homo sapiens	36,755	23-Nov-99
GB_EST8:C10137	360	C10137	C10137 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk188a1 5', mRNA sequence.	Caenorhabditis elegans	35,574	18-OCT-1999
GB_EST36:AV186952	376	AV186952	AV186952 Yuji Kohara, unpublished cDNA:Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk508a5 5', mRNA sequence.	Caenorhabditis elegans	36,702	22-Jul-99
GB_EST16:C48235	383	C48235	C48235 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk459h10 5', mRNA sequence.	Caenorhabditis elegans	42,440	18-OCT-1999
GB_BA1:MTGY227	35946	Z77724	Mycobacterium tuberculosis H37Rv complete genome, segment 114/162.	Mycobacterium tuberculosis	39,822	17-Jun-98
GB_BA1:MTV023	47852	AL022022	Mycobacterium tuberculosis H37Rv complete genome, segment 148/162.	Mycobacterium tuberculosis	39,841	17-Jun-98

**TABLE 4: ALIGNMENT RESULTS**

rx00423	576	GB_PR3:AC003091 GB_BA1:AP000063 GB_IN1:LDHSP100 GB_GSS11:AQ274393	137817 185300 7726 572	AC003091 AP000063 Z94053 AQ274393	Human BAC clone RG326C04 from 7p21, complete sequence. Aeropyrum permix genomic DNA, section 6/7. L.donovani hsp100 gene. nbxb0035G12r CUGI Rice BAC Library <i>Oryza sativa</i> genomic clone nbxb0035G12r, genomic survey sequence.	Homo sapiens Aeropyrum permix Leishmania donovani <i>Oryza sativa</i>	36,653 38,908 39,646 35,714	6-Nov-97 22-Jun-99 28-Apr-97 3-Nov-98
rx00424	594	GB_BA1:AP000063 GB_BA1:AP000063 GB_IN1:LDHSP100 GB_EST30:AV021214	185300 185300 7726 281	AP000063 AP000063 Z94053 AV021214	Aeropyrum permix genomic DNA, section 6/7. Aeropyrum permix genomic DNA, section 6/7. L.donovani hsp100 gene. AV021214 Mus musculus 18-day embryo C57BL/6J Mus musculus cDNA clone 1190021P08, mRNA sequence.	Aeropyrum permix Aeropyrum permix Leishmania donovani Mus musculus	38,225 36,735 35,125 35,849	22-Jun-99 22-Jun-99 28-Apr-97 28-Aug-99
rx00425	348	GB_HTG3:AC009278 GB_HTG3:AC009278 GB_BA2:AF127082	164119 164119 10847	AC009278 AC009278 AF127082	Homo sapiens clone 44_J_4, *** SEQUENCING IN PROGRESS ***; 14 unordered pieces. Homo sapiens clone 44_J_4, *** SEQUENCING IN PROGRESS ***; 14 unordered pieces. Myxococcus xanthus ATP-dependent protease proteolytic subunit C1pP (c1pP), ATP-dependent protease ATPase subunit C1pX (c1pX), prolyl endopeptidase precursor Pep (pep), ATP-dependent protease LonV(lonV), oligopeptidase permease homolog OppA (oppA), oligopeptidase permease homolog OppB (oppB), and oligopeptidase permease homolog OppC (oppC) genes, complete cds.	Homo sapiens Homo sapiens Myxococcus xanthus	36,705 36,705 40,995	12-Aug-99 12-Aug-99 18-MAY-1999
rx00428	756	GB_PL1:AB017080 GB_EST27:AI442425	653 541	AB017080 AI442425	Porphyrin sp. DNA, internal transcribed spacer 1 (ITS1). sa28g05.y1 Gm-c1004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1004-465 5' similar to SW:ND01_RABIT Q28615 RENAL SODIUM/DICARBOXYLATE COTRANSPORTER, mRNA sequence.	Porphyrin sp. Glycine max	38,491 40,000	10-Sep-99 01-DEC-1999
rx00429	525	GB_EST3:R33129 GB_PH:AF115103	440 40739	R33129 AF115103	yh81c08.s1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:136142 3' similar to gb:X53742_ma1 FIBULIN 1, ISOFORM B PRECURSOR (HUMAN); mRNA sequence. Streptococcus thermophilus bacteriophage Sf21, complete genome.	Homo sapiens Streptococcus thermophilus bacteriophage Sf21	37,216 36,069	28-Apr-95 18-Jul-99
rx00430	534	GB_BA1:MSGY126 GB_BA1:MTY13D12	37164 37085	AD000012 Z80343	Streptococcus thermophilus bacteriophage Sf19, complete genome. Mycobacterium tuberculosis sequence from clone y126. Mycobacterium tuberculosis H37Rv complete genome; segment 156/162.	Streptococcus thermophilus bacteriophage Sf19 Mycobacterium tuberculosis Mycobacterium tuberculosis	55,491 55,491	10-DEC-1996 17-Jun-98
rx00433	648	GB_BA1:MSGB971CS GB_PR2:AP000073 GB_IN1:CELF29G9 GB_BA1:MSGY414A	37566 100000 42751 40121	L78821 AP000073 AF016440 AD000007	Mycobacterium leprae cosmid B971 DNA sequence. Homo sapiens genomic DNA, chromosome 8p11.2, senescence gene region, section 9/19, complete sequence. Caenorhabditis elegans cosmid F29G9. Mycobacterium tuberculosis sequence from clone y414a.	Mycobacterium leprae Homo sapiens Caenorhabditis elegans Mycobacterium tuberculosis	36,905 38,043 35,474 36,809	15-Jun-96 20-Nov-99 7-Aug-97 03-DEC-1996
rx00451	615	GB_EST26:AI389267 GB_EST37:AI945493	643 574	AI389267 AI945493	GH20396.5prime GH <i>Drosophila melanogaster</i> head pOT2 <i>Drosophila melanogaster</i> cDNA clone GH20396 5prime, mRNA sequence. bs13e05.y1 <i>Drosophila melanogaster</i> adult testis library <i>Drosophila melanogaster</i> cDNA clone bs13e05 5' mRNA sequence.	<i>Drosophila melanogaster</i> <i>Drosophila melanogaster</i>	41,085 44,040	28-Jan-99 17-Aug-99
rx00455	873	GB_GSS11:AQ288118 GB_IN1:DMOVO GB_EST14:AA390588	630 6655 513	AQ288118 X59772 AA390588	nbxb003218r CUGI Rice BAC Library <i>Oryza sativa</i> genomic clone nbxb003218r, genomic survey sequence. D.melanogaster ovo gene required for female germ line development. LD09657.5prime LD <i>Drosophila melanogaster</i> embryo BlueScript <i>Drosophila melanogaster</i> cDNA clone LD09657 5prime, mRNA sequence.	<i>Oryza sativa</i> <i>Drosophila melanogaster</i> <i>Drosophila melanogaster</i>	37,885 35,104 39,759	03-DEC-1998 24-Feb-99 28-Nov-98

TABLE 4: ALIGNMENT RESULTS

rx00457	1203	GB_EST19:AA081874	621	AA081874	GM03519.5prime GM Drosophila melanogaster ovary BlueScript Drosophila melanogaster cDNA clone GM03519 5prime similar to U11383: ovo FBgn0003028 PID:g520527 SWISS-PROT:P51521, mRNA sequence.	Drosophila melanogaster	35,437	25-Nov-98
		GB_GSS8:AA000125	398	AQ000125	CIT-HSP-2282P3.TF CIT-HSP Homo sapiens genomic clone 2282P3, genomic survey sequence.	Homo sapiens	41,730	26-Jun-98
		GB_IN1:DROADDLIKE	4209	L14330	Drosophila melanogaster adducin-like protein, complete cds.	Drosophila melanogaster	37,795	11-Jun-93
		GB_IN1:DROHTSC	3922	L05016	Drosophila melanogaster hu-li tai shao (hls) mRNA, complete cds.	Drosophila melanogaster	37,081	26-Apr-93
rx00462	1503	GB_HTG3:AC009210	103814	AC009210	Drosophila melanogaster chromosome 2 clone BACR01106 (D1054) RPCI-98 01.1.6 map 55D-55D strain y, cn bw sp, *** SEQUENCING IN PROGRESS *** 86 unordered pieces.	Drosophila melanogaster	33,356	20-Aug-99
		GB_HTG3:AC009210	103814	AC009210	Drosophila melanogaster chromosome 2 clone BACR01106 (D1054) RPCI-98 01.1.6 map 55D-55D strain y, cn bw sp, *** SEQUENCING IN PROGRESS *** 86 unordered pieces.	Drosophila melanogaster	33,356	20-Aug-99
		GB_BA1:SC09	37730	AL049841	Streptomyces coelicolor cosmid E9.	Streptomyces coelicolor	37,308	19-MAY-1999
rx00463	945	GB_BA2:AF052652	2096	AF052652	Corynebacterium glutamicum homoserine O-acetyltransferase (metA) gene, complete cds.	Corynebacterium glutamicum	99,481	19-MAR-1998
		GB_GSS12:AQ047770	500	AQ047770	HS_5089_B1_F03_TTA RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=645 Col=5 Row=L, genomic survey sequence.	Homo sapiens	40,081	17-MAR-1999
		GB_GSS15:AQ596209	358	AQ596209	HS_5482_A2_H10_SPEE RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1058 Col=20 Row=O, genomic survey sequence.	Homo sapiens	40,000	8-Jun-99
rx00468	942	GB_EST32:AI763196	341	AI763196	wf65h04.x1 NCI CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2398231 3', mRNA sequence.	Homo sapiens	38,824	24-Jun-99
		GB_EST17:AA652964	329	AA652964	ns62e02.s1 NCI CGAP_P122 Homo sapiens cDNA clone IMAGE:1188218 3', mRNA sequence.	Homo sapiens	41,104	13-Nov-97
		GB_EST20:AA864303	411	AA864303	oh54d02.s1 NCI CGAP_G04 Homo sapiens cDNA clone IMAGE:1470435 3', mRNA sequence.	Homo sapiens	41,422	13-MAY-1998
rx00469	1299	GB_PR4:AC005988	173126	AC005988	Homo sapiens chromosome 17, clone hRPK.299_G_24, complete sequence.	Homo sapiens	34,699	15-Jan-99
		GB_PR4:AC005988	173126	AC005988	Homo sapiens chromosome 17, clone hRPK.299_G_24, complete sequence.	Homo sapiens	35,725	15-Jan-99
		GB_HTG3:AC009116	186292	AC009116	Homo sapiens chromosome 16 clone RPCI-11_477D3, *** SEQUENCING IN PROGRESS *** 47 unordered pieces.	Homo sapiens	36,222	3-Aug-99
rx00472	942	GB_HTG3:AC007882	214882	AC007882	Homo sapiens clone NH0499D05, *** SEQUENCING IN PROGRESS *** 2 unordered pieces.	Homo sapiens	38,245	8-Sep-99
		GB_HTG3:AC007882	214882	AC007882	Homo sapiens clone NH0499D05, *** SEQUENCING IN PROGRESS *** 2 unordered pieces.	Homo sapiens	38,245	8-Sep-99
		GB_PR2:HUA002038	161973	AC002038	Homo sapiens chromosome 2 clone 101B6 map 2p11, complete sequence.	Homo sapiens	37,961	30-Jun-97
rx00473	912	GB_HTG3:AC011445	144370	AC011445	Homo sapiens chromosome 19 clone CIT-HSPC_246B18, *** SEQUENCING IN PROGRESS *** 31 unordered pieces.	Homo sapiens	38,470	07-OCT-1999
		GB_HTG3:AC011445	144370	AC011445	Homo sapiens chromosome 19 clone CIT-HSPC_246B18, *** SEQUENCING IN PROGRESS *** 31 unordered pieces.	Homo sapiens	38,470	07-OCT-1999
		GB_RO:AB026437	2097	AB026437	Mus musculus DNA, 5' flanking region of interleukin 12 receptor beta 1.	Mus musculus	40,000	02-OCT-1999
rx00474	1701	GB_GSS10:AQ223838	543	AQ223838	HS_2218_A1_H03_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2218 Col=5 Row=O, genomic survey sequence.	Homo sapiens	40,189	20-Sep-98
		GB_GSS10:AQ223838	543	AQ223838	HS_2218_A1_H03_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2218 Col=5 Row=O, genomic survey sequence.	Homo sapiens	37,944	20-Sep-98
rx00475	783	GB_PL2:GMAKHSDDH2	10535	AF049708	Glycine max aspartokinase-homoserine dehydrogenase (AK-HSDH) gene, partial cds.	Glycine max	36,446	7-Jul-99
		GB_EST14:AA386651	351	AA386651	vb54b04.r1 Ko mouse embryo 11 5dpc Mus musculus cDNA clone IMAGE:760783 5', mRNA sequence.	Mus musculus	41,311	23-Apr-97
		GB_EST14:AA386603	498	AA386603	vb53c02.r1 Ko mouse embryo 11 5dpc Mus musculus cDNA clone IMAGE:760706 5' similar to TR:G56689 G56689 DIMETHYLGLYCINE DEHYDROGENASE, mRNA sequence.	Mus musculus	40,644	23-Apr-97
rx00476	984	GB_GSS5:AQ770769	554	AQ770769	HS_5357_B2_H01_TTA RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=933 Col=2 Row=P, genomic survey sequence.	Homo sapiens	35,560	28-Jul-99
		GB_IN1:CEM162	39977	Z82278	Caenorhabditis elegans cosmid M162, complete sequence.	Caenorhabditis elegans	34,224	19-Nov-99
		GB_IN1:CEM162	39977	Z82278	Caenorhabditis elegans cosmid M162, complete sequence.	Caenorhabditis elegans	33,777	19-Nov-99
rx00481	708	GB_PR4:AC005013	195910	AC005013	Homo sapiens BAC clone GS165L15 from 7p15, complete sequence.	Homo sapiens	35,755	28-Nov-98
		GB_HTG1:PFMAL4P4	224448	AL035477	Plasmodium falciparum chromosome 4 strain 3D7, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Plasmodium falciparum	37,213	11-Aug-99

TABLE 4: ALIGNMENT RESULTS

GB_HTG1:PFMALP4	224448	AL035477	Plasmodium falciparum chromosome 4 strain 3D7, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Plasmodium falciparum	37,213	11-Aug-99
GB_EST30:AV017239	238	AV017239	AV017239 Mus musculus 18-day embryo C57BL/6J Mus musculus cDNA clone 1110069G23, mRNA sequence.	Mus musculus	39,916	28-Aug-99
GB_EST33:AV093875	254	AV093875	AV093875 Mus musculus C57BL/6J ES cell Mus musculus cDNA clone 2400006D21, mRNA sequence.	Mus musculus	38,189	22-Nov-99
GB_EST33:AV084536	287	AV084536	AV084536 Mus musculus tongue C57BL/6J adult Mus musculus cDNA clone 2310007K01, mRNA sequence.	Mus musculus	37,282	25-Jun-99
GB_RO:MUSPVPR2	1089	AF098867	Mus sp. 129SV V3V1b vasopressin receptor gene, exon 2 and complete cds.	Mus musculus	38,163	28-Apr-99
GB_GSS9:AQ166448	407	AQ166448	HS_3137_B2_A06 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3137 Col=12 Row=B, genomic survey sequence.	Homo sapiens	40,250	16-OCT-1998
GB_GSS15:AQ614261	505	AQ614261	HS_5123_B1_F11_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=899 Col=21 Row=L, genomic survey sequence.	Homo sapiens	37,905	15-Jun-99
GB_EST15:AA463205	282	AA463205	zx71c06.st Soares_t0tal_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:796906 3', mRNA sequence.	Homo sapiens	39,502	10-Jun-97
GB_BA1:SLINC	36270	X79146	S.lincolnensis (78-11) Lincomycin production genes.	Streptomyces lincolnensis	37,278	15-MAY-1996
GB_GSS3:B10984	646	B10984	F2218-SP6 IGF Arabidopsis thaliana genomic clone F2218, genomic survey sequence.	Arabidopsis thaliana	39,205	14-MAY-1997
GB_HTG1:CEY87G2	330612	AL022597	Caenorhabditis elegans chromosome 1 clone Y87G2, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	37,405	26-OCT-1999
GB_HTG1:CEY87G2	330612	AL022597	Caenorhabditis elegans chromosome 1 clone Y87G2, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	37,405	26-OCT-1999
GB_HTG1:CEY6B3	253516	Z92865	Caenorhabditis elegans chromosome 1 clone Y6B3, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	38,213	30-Jul-98
GB_BA1:SC6G4	41055	AL031317	Streptomyces coelicolor cosmid 6G4.	Streptomyces coelicolor	61,649	20-Aug-98
GB_BA2:U00015	42325	U00015	Mycobacterium leprae cosmid B1620.	Mycobacterium leprae	38,567	01-MAR-1994
GB_BA1:U00020	36947	U00020	Mycobacterium leprae cosmid B229.	Mycobacterium leprae	38,567	01-MAR-1994
GB_VI:TVRNAP	6404	X68414	Toscana Virus genomic RNA for RNA-dependent RNA polymerase.	Toscana virus	35,428	27-OCT-1992
GB_IN2:DMU09808	9239	U09808	Drosophila melanogaster twisted gastrulation (tsg) and serine protease (gd) genes, complete cds.	Drosophila melanogaster	36,837	25-Jun-98
GB_IN2:DMU09808	9239	U09808	Drosophila melanogaster twisted gastrulation (tsg) and serine protease (gd) genes, complete cds.	Drosophila melanogaster	37,782	25-Jun-98
GB_BA1:MTCY159	33818	Z83863	Mycobacterium tuberculosis H37Rv complete genome; segment 111/162.	Mycobacterium tuberculosis	36,961	17-Jun-98
GB_PR1:HUMHM145	2214	D10925	Human mRNA for HM145.	Homo sapiens	36,066	3-Feb-99
GB_EST14:AA415083	332	AA415083	Mg0017 RCW Lambda Zap Express Library Pyricularia grisea cDNA clone RCW17 5', mRNA sequence.	Pyricularia grisea	40,181	09-DEC-1999
GB_PAT:192047	551	I92047	Sequence 14 from patent US 5726299.	Unknown.	46,250	01-DEC-1998
GB_PAT:178759	549	I78759	Sequence 15 from patent US 5693781.	Unknown.	44,813	3-Apr-98
GB_PAT:192048	549	I92048	Sequence 15 from patent US 5726299.	Unknown.	44,813	01-DEC-1998
GB_PR2:HS1063B2	114596	AL035683	Human DNA sequence from clone 1063B2 on chromosome 20q13.1-13.2. Contains the 3' part of the gene for Beta-1,4-galactosyltransferase, ESTs and GSSs, complete sequence.	Homo sapiens	36,449	23-Nov-99
GB_HTG2:AC007225	218892	AC007225	Homo sapiens chromosome 16 clone 480G7, *** SEQUENCING IN PROGRESS ***; 38 unordered pieces.	Homo sapiens	36,646	6-Apr-99
GB_HTG2:AC007225	218892	AC007225	Homo sapiens chromosome 16 clone 480G7, *** SEQUENCING IN PROGRESS ***; 38 unordered pieces.	Homo sapiens	36,646	6-Apr-99
GB_GSS4:AQ707590	499	AQ707590	HS_5560_B1_H02_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1136 Col=3 Row=P, genomic survey sequence.	Homo sapiens	37,275	7-Jul-99
GB_GSS4:AQ707590	499	AQ707590	HS_5560_B1_H02_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1136 Col=3 Row=P, genomic survey sequence.	Homo sapiens	37,275	7-Jul-99

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

rx00515	825	GB_BA1:CGICD	3595	X71489	C.glutamicum icd gene for monomeric isocitrate dehydrogenase.	Corynebacterium glutamicum	100,000	17-Feb-95
		GB_BA1:CGICD	3595	X71489	C.glutamicum icd gene for monomeric isocitrate dehydrogenase.	Corynebacterium glutamicum	38,150	17-Feb-95
		GB_GSS13:AQ451896	509	AQ451896	HS_5184_B1_C03_SP6E RPCL1-11 Human Male BAC Library Homo sapiens genomic clone Plate=760 Col=5 Row=F, genomic survey sequence.	Homo sapiens	36,638	21-Apr-99
rx00519	2337	GB_BA1:CGICD	3595	X71489	C.glutamicum icd gene for monomeric isocitrate dehydrogenase.	Corynebacterium glutamicum	100,000	17-Feb-95
		GB_BA2:AF127018	2560	AF127018	Streptomyces coelicolor isocitrate dehydrogenase (idh) gene, idh-B allele, complete cds.	Streptomyces coelicolor	66,667	1-Jun-99
		GB_BA1:AVIICD	3550	D73443	Azotobacter vinelandii icd gene for isocitrate dehydrogenase, complete cds.	Azotobacter vinelandii	63,652	4-Feb-99
rx00527	1887	GB_PAT:192049	2248	192049	Sequence 16 from patent US 5726299.	Unknown.	39,250	01-DEC-1998
		GB_PAT:192053	2213	192053	Sequence 20 from patent US 5726299.	Unknown.	70,635	01-DEC-1998
		GB_BA1:MTCY98	31225	Z83860	Mycobacterium tuberculosis H37Rv complete genome, segment 103/162.	Mycobacterium tuberculosis	37,741	17-Jun-98
rx00528	1212	GB_BA1:MSGY219	38721	AD000013	Mycobacterium tuberculosis sequence from clone y219.	Mycobacterium tuberculosis	68,672	10-DEC-1996
		GB_BA1:MTCY21D4	20760	Z80775	Mycobacterium tuberculosis H37Rv complete genome, segment 3/262.	Mycobacterium tuberculosis	39,762	24-Jun-99
rx00529	666	GB_BA1:SCH24	41625	AL049826	Streptomyces coelicolor cosmid H24.	Streptomyces coelicolor	40,411	11-MAY-1999
		GB_PR2:HSAC000109	41122	AC000109	Human Cosmid g0771a222 from 7q31.3, complete sequence.	Homo sapiens	37,462	11-Sep-97
		GB_PR2:HSAC000110	45508	AC000110	Human Cosmid g0771a233, complete sequence.	Homo sapiens	37,462	30-Jan-97
		GB_PR2:HSAC000109	41122	AC000109	Human Cosmid g0771a222 from 7q31.3, complete sequence.	Homo sapiens	39,724	11-Sep-97
rx00530	1404	GB_PR3:HS435C23	151798	Z92844	Human DNA sequence from PAC 435C23 on chromosome X. Contains ESTs.	Homo sapiens	36,482	23-Nov-99
		GB_PR3:HS435C23	151798	Z92844	Human DNA sequence from PAC 435C23 on chromosome X. Contains ESTs.	Homo sapiens	37,918	23-Nov-99
		GB_PL1:YSCADE3	4883	M12878	Saccharomyces cerevisiae C-1-tetrahydrofolate synthase (ADE3) gene, complete cds.	Saccharomyces cerevisiae	37,034	11-MAY-1995
rx00535	840	GB_BA1:CGLEUA	3492	X70959	C.glutamicum gene leuA for isopropylmalate synthase.	Corynebacterium glutamicum	100,000	10-Feb-99
		GB_BA1:CORASKD	2957	L16848	Corynebacterium flavum aspartokinase (ask), and aspartate-semialdehyde dehydrogenase (asd) genes, complete cds.	Corynebacterium flavescens	43,750	11-Jun-93
		GB_GSS10:AQ193141	515	AQ193141	HS_3060_B1_F11_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3060 Col=21 Row=L, genomic survey sequence.	Homo sapiens	44,773	4-Nov-98
rx00540	366	GB_PAT:192052	2115	192052	Sequence 19 from patent US 5726299.	Unknown.	74,795	01-DEC-1998
		GB_HTG2:AC008095	126322	AC008095	Drosophila melanogaster chromosome 2 clone BACR11H16 (D932) RPCL-98 11 H.16 map 52A-52A strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** 95 unordered pieces.	Drosophila melanogaster	41,899	2-Aug-99
		GB_HTG2:AC008095	126322	AC008095	Drosophila melanogaster chromosome 2 clone BACR11H16 (D932) RPCL-98 11 H.16 map 52A-52A strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** 95 unordered pieces.	Drosophila melanogaster	41,899	2-Aug-99
rx00547	1521	GB_BA1:MSGY219	38721	AD000013	Mycobacterium tuberculosis sequence from clone y219.	Mycobacterium tuberculosis	36,910	10-DEC-1996
		GB_BA1:MTCY21D4	20760	Z80775	Mycobacterium tuberculosis H37Rv complete genome, segment 3/262.	Mycobacterium tuberculosis	51,125	24-Jun-99
		GB_EST27:A1415174	292	A1415174	mc05c02.x1 Soares mouse p3NMF19.5 Mus musculus cDNA clone IMAGE:338018 3' mRNA sequence.	Mus musculus	39,384	9-Feb-99
rx00549	1797	GB_PL2:ATAC007135	27313	AC007135	Arabidopsis thaliana chromosome II BAC F9C22 genomic sequence, complete sequence.	Arabidopsis thaliana	35,584	26-MAY-1999
		GB_PL2:ATAC006921	76042	AC006921	Arabidopsis thaliana chromosome II BAC F2H17 genomic sequence, complete sequence.	Arabidopsis thaliana	36,581	23-MAR-1999
		GB_PL2:ATAC007135	27313	AC007135	Arabidopsis thaliana chromosome II BAC F9C22 genomic sequence, complete sequence.	Arabidopsis thaliana	35,827	26-MAY-1999
rx00550								
rx00552	1059	GB_BA1:D90742	19201	D90742	Escherichia coli genomic DNA. (23.8 - 24.2 min).	Escherichia coli	46,072	7-Feb-99
		GB_BA1:ECHTRB	3129	X61000	E. coli K12 HfrB gene.	Escherichia coli	39,164	30-Jun-93

TABLE 4: ALIGNMENT RESULTS

rx00553	444	GB_BA2:AE000207	11148	AE000207	Escherichia coli K-12 MG1655 section 97 of 400 of the complete genome.	46,072	12-Nov-98
		GB_EST18:AB009093	479	AB009093	Chlamydomonas W80 lambda ZAP II Chlamydomonas sp. cDNA similar to photosystem II 10 kDa protein. mRNA sequence.	41,808	05-DEC-1997
		GB_EST30:AI640954	641	AI640954	AEMTAP02 Aedes aegypti MT pSPORT Library Aedes aegypti cDNA clone AP02 5'. mRNA sequence.	38,991	28-Apr-99
		GB_GSS13:AQ467517	206	AQ467517	HS_5219_A2_F02_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=795 Col=4 Row=K. genomic survey sequence.	45,255	23-Apr-99
rx00554	594	GB_EST6:W04418	364	W04418	zaf3c06.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:295306 5'. mRNA sequence.	39,688	22-Apr-96
		GB_EST35:AL041829	564	AL041829	DKFZp434C0318_s1 434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434C0318 3'. mRNA sequence.	40,433	29-Sep-99
		GB_EST35:AL041828	386	AL041828	DKFZp434C0318_r1 434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434C0318 5'. mRNA sequence.	39,688	29-Sep-99
rx00555							
rx00560	498	GB_BA1:MTCY7H7A	10451	Z95618	Mycobacterium tuberculosis H37Rv complete genome; segment 39/162.	52,727	17-Jun-98
		GB_BA1:BAPURF	1885	X91252	B.ammoniaenes purf gene.	61,092	5-Jun-97
		GB_PL1:YSCMET10A	3650	L26504	Saccharomyces carlsbergensis assimilatory sulfite reductase (MET10) gene, complete cds.	41,273	7-Feb-95
rx00563	2762	GB_BA1:BAFASAA	10549	X64795	B.ammoniaenes FAS gene.	66,910	14-OCT-1997
		GB_BA1:MTCY159	33818	Z83863	Mycobacterium tuberculosis H37Rv complete genome; segment 111/162.	40,066	17-Jun-98
rx00564	528	GB_BA1:MBU36763	8391	U36763	Mycobacterium bovis fatty acid synthase gene, complete cds.	61,178	15-Jul-96
		GB_PR3:HS833B7	86574	AL008637	Human DNA sequence from clone 833B7 on chromosome 22q12.3-13.2 Contains genes for NCF4 (P40PHOX) protein cytokine receptor common beta chain precursor CSF2RB (partial), ESTs, CA repeat, STS, GSS, complete sequence.	39,015	23-Nov-99
		GB_HTG3:AC008543	278334	AC008543	Homo sapiens chromosome 19 clone CIT-HSPC_499B15. *** SEQUENCING IN PROGRESS ***.	36,328	2-Sep-99
		GB_HTG3:AC008543	278334	AC008543	134 unordered pieces.	36,328	2-Sep-99
rx00573							
rx00574	1002	GB_GSS11:AQ301816	481	AQ301816	HS_3174_A1_B04_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3174 Col=7 Row=C. genomic survey sequence.	43,137	16-DEC-1998
		GB_PR3:AC004537	88872	AC004537	Homo sapiens PAC clone DJ0872F07 from 7q31. complete sequence.	34,712	9-Apr-98
		GB_EST29:AI563059	339	AI563059	EST00183 watermelon lambda zap library Citrullus lanatus cDNA clone WMLS355 5' similar to unknown protein. mRNA sequence.	37,758	26-MAR-1999
rx00576	795	GB_EST37:AI947508	533	AI947508	603022E09.x1 603 - stressed root cDNA library from Wang/Bohmer lab Zea mays cDNA. mRNA sequence.	38,728	19-Aug-99
		GB_GSS11:AQ296770	347	AQ296770	HS_3087_A2_B12_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3087 Col=24 Row=C. genomic survey sequence.	40,058	15-DEC-1998
		GB_GSS13:AQ503769	589	AQ503769	RPCI-11-282O13.TV RPCI-11 Homo sapiens genomic clone RPCI-11-282O13, genomic survey sequence.	37,993	29-Apr-99
rx00577	471	GB_GSS14:AQ533027	638	AQ533027	RPCI-11-351M24.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-351M24, genomic survey sequence.	34,944	18-MAY-1999



**TABLE 4: ALIGNMENT RESULTS**

rx00582	642	GB_PR3:HS440B3	28047	AL022331	Homo sapiens DNA sequence from clone 440B3 on chromosome 22q12.1-3 Contains a pseudogene similar to 60S Ribosomal protein L17. Contains ESTs and an STS (genomic marker D22S1176), complete sequence.	Homo sapiens	33,626	23-Nov-99
		GB_PL2:ZEU19267	1230	U19267	Zinnia elegans cysteine proteinase mRNA, complete cds.	Zinnia elegans	35,456	26-Aug-96
rx00582	642	GB_BA1:CORAHPS	2570	L07603	Corynebacterium glutamicum 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase gene, complete cds.	Corynebacterium glutamicum	98,901	26-Apr-93
		GB_PR3:AC005389	120359	AC005389	Homo sapiens chromosome 17, clone HRPK.601.N.13, complete sequence.	Homo sapiens	38,315	14-Aug-98
		GB_HTG6:AC008002	126629	AC008002	Drosophila melanogaster chromosome 2 clone BACR48E08 (D843) RPCI-98 48.E.8 map 21D-21E strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 85 unordered pieces.	Drosophila melanogaster	32,437	07-DEC-1999
rx00585	441	GB_BA1:CORAHPS	2570	L07603	Corynebacterium glutamicum 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase gene, complete cds.	Corynebacterium glutamicum	98,039	26-Apr-93
		GB_PR4:AC004970	149951	AC004970	Homo sapiens BAC clone DJ1122F04 from 7q11.23-q21.2, complete sequence.	Homo sapiens	39,900	27-Aug-99
		GB_PR2:HS102G20	99207	Z99127	Human DNA sequence from PAC 102G20 on chromosome 1q24-q25. Contains ESTs, STSs and a predicted CpG island.	Homo sapiens	41,509	23-Nov-99
rx00586	1005	GB_BA1:MTV017	67200	AL021897	Mycobacterium tuberculosis H37Rv complete genome, segment 48/162.	Mycobacterium tuberculosis	56,219	24-Jun-99
		GB_BA1:MLU15183	36800	U15183	Mycobacterium leprae cosmid B1740.	Mycobacterium leprae	55,622	09-MAR-1995
		GB_BA1:MTV017	67200	AL021897	Mycobacterium tuberculosis H37Rv complete genome, segment 48/162.	Mycobacterium tuberculosis	37,838	24-Jun-99
rx00587	459	GB_EST37:AI993539	514	AI993539	701496589 A. thaliana, Ohio State clone set Arabidopsis thaliana cDNA clone 701496589, mRNA sequence.	Arabidopsis thaliana	40,153	8-Sep-99
		GB_GSS10:AQ224941	511	AQ224941	HS_2009_B1_B06_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2009 Col=11 Row=D, genomic survey sequence.	Homo sapiens	43,750	20-Sep-98
rx00589	573	GB_EST23:AI099719	475	AI099719	33872 Lambda-PRL2 Arabidopsis thaliana cDNA clone 120M10XP 3', mRNA sequence.	Arabidopsis thaliana	36,752	21-Aug-98
		GB_PL2:ATAC003673	70575	AC003673	Arabidopsis thaliana chromosome II BAC F19F24 genomic sequence, complete sequence.	Arabidopsis thaliana	39,785	1-Apr-98
		GB_PR3:HS427A4	149466	Z98049	Human DNA sequence from PAC 427A4 on chromosome 6q26-q27. Contains ribosomal protein S6 kinase, RSK3, ESTs, CpG island.	Homo sapiens	35,145	23-Nov-99
rx00595		GB_PL2:ATAC005724	86671	AC005724	Arabidopsis thaliana chromosome II P1 MSF3 genomic sequence, complete sequence.	Arabidopsis thaliana	39,785	24-Jan-99
rx00597	393	GB_PR3:AC004659	129577	AC004659	Homo sapiens chromosome 19, CIT-HSP-87m17 BAC clone, complete sequence.	Homo sapiens	39,459	02-MAY-1998
		GB_GSS14:AQ575039	927	AQ575039	nbx0086L01r CUGI Rice BAC Library Oryza sativa genomic clone nbx0086L01r, genomic survey sequence.	Oryza sativa	37,786	2-Jun-99
rx00598		GB_BA1:NMRRNA	5209	X72495	N.magadii rRNA operon.	Natrialba magadii	39,788	10-Feb-95
rx00601	414	GB_BA2:AF175719	1368	AF175719	Porphyromonas gingivalis strain W50 immunoreactive 51kD antigen PG52 gene, complete cds.	Porphyromonas gingivalis	35,331	23-Aug-99
		GB_GSS9:AQ140775	464	AQ140775	HS_3128_A1_B11_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3128 Col=21 Row=C, genomic survey sequence.	Homo sapiens	40,431	24-Sep-98
		GB_EST8:AA018824	542	AA018824	ze57e09.s1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:363112 3', mRNA sequence.	Homo sapiens	40,590	30-Jan-97
rx00602	876	GB_EST30:AI642687	479	AI642687	w02h03.x1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone IMAGE:1230773 3', mRNA sequence.	Mus musculus	40,042	29-Apr-99
		GB_EST20:AA879989	412	AA879989	w03a05.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone IMAGE:1230800 5', mRNA sequence.	Mus musculus	39,948	26-MAR-1998

TABLE 4: ALIGNMENT RESULTS

rx00604	414	GB_EST28:AL481047	438	AL481047	v91a05.x1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone IMAGE:851120.3; mRNA sequence.	Mus musculus	38,128	09-MAR-1999
		GB_HTG2:AC008205	131658	AC008205	Drosophila melanogaster chromosome 3 clone BACR33F18 (D764) RPCI-98 33.F.18 map 96A-96B strain y, cn bw sp. *** SEQUENCING IN PROGRESS ***; 118 unordered pieces.	Drosophila melanogaster	33,907	2-Aug-99
		GB_HTG2:AC008205	131658	AC008205	Drosophila melanogaster chromosome 3 clone BACR33F18 (D764) RPCI-98 33.F.18 map 96A-96B strain y, cn bw sp. *** SEQUENCING IN PROGRESS ***; 118 unordered pieces.	Drosophila melanogaster	33,907	2-Aug-99
		GB_IN1:DMBR7C10	56820	AL121804	Drosophila melanogaster clone BACR7C10.	Drosophila melanogaster	44,267	10-OCT-1999
rx00610	987	GB_BA1:SME242575	1403	AJ242575	Sinorhizobium meliloti partial ox1 and dehydrogenase genes, isolate Ipu119.	Sinorhizobium meliloti	41,760	26-MAY-1999
		GB_PR3:AC004655	134929	AC004655	Homo sapiens Xp22-140-141 BAC GSHB-12865 (Genome Systems Human BAC library) complete sequence.	Homo sapiens	38,422	17-Sep-98
		GB_PR3:HSS98F2	99886	AL021579	Human DNA sequence from clone 598F2 on chromosome 1q23.1-24.3 Contains ESTs, STS and GSS, complete sequence.	Homo sapiens	38,351	23-Nov-99
		GB_HTG2:AC007108	190000	AC007108	Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***; 24 unordered pieces.	Homo sapiens	17,451	17-MAR-1999
		GB_HTG2:AC007108	190000	AC007108	Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***; 24 unordered pieces.	Homo sapiens	17,451	17-MAR-1999
		GB_HTG4:AC010893	176178	AC010893	Homo sapiens chromosome unknown clone NH0480A20, WORKING DRAFT SEQUENCE. in unordered pieces.	Homo sapiens	35,819	29-OCT-1999
		GB_IN2:AC004361	87747	AC004361	Drosophila melanogaster DNA sequence (P1 DS07851 (D49)), complete sequence.	Drosophila melanogaster	35,081	29-MAY-1998
		GB_PL2:AC006268	105420	AC006268	Arabidopsis thaliana BAC T24G23 from chromosome IV near 21 cM, complete sequence.	Arabidopsis thaliana	43,682	1-Jan-99
		GB_BA1:MLCB596	38426	AL035472	Mycobacterium leprae cosmid B596.	Mycobacterium leprae	35,026	27-Aug-99
rx00614	1038	GB_BA1:MTV025	121125	AL022121	Mycobacterium tuberculosis H37Rv complete genome, segment 155/162.	Mycobacterium tuberculosis	53,061	24-Jun-99
		GB_BA1:SCH66	9153	AL049731	Streptomyces coelicolor cosmid H66.	Streptomyces coelicolor	52,817	29-Apr-99
		GB_EST14:AA446728	411	AA446728	zw84f03.r1 Soares, total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:783677.5; mRNA sequence.	Homo sapiens	36,548	3-Jun-97
rx00616								
rx00617	351	GB_PL1:AB009030	2589	AB009030	Panax ginseng OSCPNY1 mRNA for beta-Amyrin Synthase, complete cds.	Panax ginseng	39,048	03-OCT-1998
		GB_PR3:HS905G11	122469	AL035045	Human DNA sequence from clone 905G11 on chromosome 20p11.2-12.1. Contains STSs, GSSs and genomic marker D20S182, complete sequence.	Homo sapiens	39,255	23-Nov-99
		GB_RO:MMT1CPS	8147	X15147	Mouse T1a region T1c pseudogene for class I antigen major histocompatibility complex.	Mus musculus	36,311	19-Feb-90
rx00628	531	GB_HTG5:AC010674	220575	AC010674	Homo sapiens chromosome 15 clone RP11-430B1 map 15q21, *** SEQUENCING IN PROGRESS ***; 46 ordered pieces.	Homo sapiens	37,714	5-Nov-99
		GB_HTG5:AC010674	220575	AC010674	Homo sapiens chromosome 15 clone RP11-430B1 map 15q21, *** SEQUENCING IN PROGRESS ***; 46 ordered pieces.	Homo sapiens	39,293	5-Nov-99
		GB_BA1:BRLBIOAD	2272	D14083	Brevibacterium flavum genes for 7,8-diaminopelargonic acid aminotransferase and dethiobiotin synthetase, complete cds.	Corynebacterium glutamicum	47,368	3-Feb-99
		GB_PAT:E04041	675	E04041	DNA sequence coding for dethiobiotinsynthetase.	Corynebacterium glutamicum	46,552	29-Sep-97
		GB_EST20:AA820386	453	AA820386	LD23968.5prime LD Drosophila melanogaster embryo pOT2 Drosophila melanogaster cDNA clone LD23968.5prime, mRNA sequence.	Drosophila melanogaster	45,679	25-Feb-99
rx00637	876	GB_HTG2:AC007589	134659	AC007589	Drosophila melanogaster chromosome 3 clone BACR20D10 (D667) RPCI-98 20.D.10 map 82D-82E strain y, cn bw sp. *** SEQUENCING IN PROGRESS ***; 73 unordered pieces.	Drosophila melanogaster	32,102	2-Aug-99
		GB_HTG2:AC007589	134659	AC007589	Drosophila melanogaster chromosome 3 clone BACR20D10 (D667) RPCI-98 20.D.10 map 82D-82E strain y, cn bw sp. *** SEQUENCING IN PROGRESS ***; 73 unordered pieces.	Drosophila melanogaster	32,102	2-Aug-99
		GB_HTG3:AC009212	125452	AC009212	Drosophila melanogaster chromosome 3 clone BACR01A18 (D669) RPCI-98 01.A.18 map 82E-82F strain y, cn bw sp. *** SEQUENCING IN PROGRESS ***; 119 unordered pieces.	Drosophila melanogaster	37,126	23-Aug-99
rx00646	541	GB_HTG1:AP000488	123363	AP000488	Homo sapiens chromosome 11 clone B759H8 map 11q23, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,264	13-Sep-99

TABLE 4: ALIGNMENT RESULTS

GB_HTG1:AP000488	123363	AP000488	123363	Homo sapiens chromosome 11 clone B759H8 map 11q23, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Homo sapiens	38,264	13-Sep-99
GB_HTG1:AP000488	123363	AP000488	123363	Homo sapiens chromosome 11 clone B759H8 map 11q23, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Homo sapiens	36,484	13-Sep-99
GB_GSS13:QA431426	536	AQ431426	536	HS_5140_A2_E01_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=716 Col=2 Row=1, genomic survey sequence.	Homo sapiens	35,635	31-MAR-1999
GB_OV:CHKP4HA	3149	M26217	3149	Chicken prolyl 4-hydroxylase alpha subunit gene, 3' end.	Gallus gallus	40,655	28-Apr-93
GB_HTG2:AC008271	168302	AC008271	168302	Homo sapiens clone NH0123E16, *** SEQUENCING IN PROGRESS ***. 2 unordered pieces.	Homo sapiens	37,417	31-Jul-99
GB_PR2:AC002563	136436	AC002563	136436	Human PAC clone 127H14 from 12q, complete sequence.	Homo sapiens	37,937	26-Sep-97
GB_HTG3:AC011466	165953	AC011466	165953	Homo sapiens chromosome 19 clone CIT-HSPC_453023, *** SEQUENCING IN PROGRESS ***, 74 unordered pieces.	Homo sapiens	38,179	07-OCT-1999
GB_HTG3:AC011466	165953	AC011466	165953	Homo sapiens chromosome 19 clone CIT-HSPC_453023, *** SEQUENCING IN PROGRESS ***, 74 unordered pieces.	Homo sapiens	38,179	07-OCT-1999
GB_EST1:Z34080	271	Z34080	271	ATTS3128 Grenoble-B Arabidopsis thaliana cDNA clone GBGe328 5', mRNA sequence.	Arabidopsis thaliana	40,370	6-Jun-94
GB_PR3:AC000460	113803	AC000460	113803	Homo sapiens PAC clone DJ1088D14, complete sequence.	Homo sapiens	36,150	24-MAR-1998
GB_GSS6:QA835185	571	AQ835185	571	HS_4832_A1_E02_T7A CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=4832 Col=3 Row=1, genomic survey sequence.	Homo sapiens	39,429	27-Aug-99
GB_HTG3:AC009948	172463	AC009948	172463	Homo sapiens clone NH0065L03, *** SEQUENCING IN PROGRESS ***. 2 unordered pieces.	Homo sapiens	43,164	25-Sep-99
GB_HTG3:AC009948	172463	AC009948	172463	Homo sapiens clone NH0065L03, *** SEQUENCING IN PROGRESS ***. 2 unordered pieces.	Homo sapiens	43,164	25-Sep-99
GB_GSS13:QA462899	522	AQ462899	522	HS_5212_A1_C09_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=788 Col=17 Row=E, genomic survey sequence.	Homo sapiens	42,529	23-Apr-99
GB_BA2:AF064700	3481	AF064700	3481	Rhodococcus sp. NO1-1 CprS and CprR genes, complete cds.	Rhodococcus sp. NO1-1	40,558	15-Jul-98
GB_PR3:AC005346	3849	AC005346	3849	Homo sapiens chromosome 16, cosmid clone 2H2 (LANL), complete sequence.	Homo sapiens	35,553	31-Jul-98
GB_HTG3:AC008905	129915	AC008905	129915	Homo sapiens chromosome 5 clone CITB-H1_2259114, *** SEQUENCING IN PROGRESS ***. 40 unordered pieces.	Homo sapiens	37,179	3-Aug-99
GB_PL2:AC006193	118335	AC006193	118335	Arabidopsis thaliana chromosome I BAC F13O11 genomic sequence, complete sequence.	Arabidopsis thaliana	35,513	11-Jun-99
GB_RO:MMFABPE	6593	AJ223066	6593	Mus musculus Fabpe gene.	Mus musculus	37,500	27-Jul-98
GB_PL2:AC006193	118335	AC006193	118335	Arabidopsis thaliana chromosome I BAC F13O11 genomic sequence, complete sequence.	Arabidopsis thaliana	33,552	11-Jun-99
GB_EST29:AI551960	718	AI551960	718	vi48d09.y1 Beddington mouse embryonic region Mus musculus cDNA clone IMAGE:907025 5' similar to gb:D10576 Mouse mRNA for ubiquitin activating enzyme E1 (MOUSE);, mRNA sequence.	Mus musculus	39,972	23-MAR-1999
GB_BA2:AE000633	19734	AE000633	19734	Helicobacter pylori 26695 section 111 of 134 of the complete genome.	Helicobacter pylori 26695	36,606	6-Apr-99
GB_GSS10:QA216730	529	AQ216730	529	HS_2262_A1_G06_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2262 Col=11 Row=M, genomic survey sequence.	Homo sapiens	32,703	19-Sep-98
GB_HTG3:AC009205	113482	AC009205	113482	Drosophila melanogaster chromosome 2 clone BACR04C20 (D1035) RPCI-98 04.C.20 map 36E- 37C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***. 101 unordered pieces.	Drosophila melanogaster	36,713	17-Sep-99
GB_EST25:AI259480	626	AI259480	626	LP02903.5prime LP Drosophila melanogaster larval-early pupal pOT2 Drosophila melanogaster cDNA clone LP02903 5prime, mRNA sequence.	Drosophila melanogaster	37,173	17-Nov-98
GB_HTG3:AC009205	113482	AC009205	113482	Drosophila melanogaster chromosome 2 clone BACR04C20 (D1035) RPCI-98 04.C.20 map 36E- 37C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***. 101 unordered pieces.	Drosophila melanogaster	36,713	17-Sep-99
GB_PAT:AR048317	2627	AR048317	2627	Sequence 3 from patent US 5821090.	Unknown.	39,805	29-Sep-99
GB_PAT:AA46560	2627	A46560	2627	Sequence 3 from Patent WO9526406.	Eremothecium gossypii	39,805	07-MAR-1997
GB_VI:HEHCMVCG	229354	X17403	229354	Human cytomegalovirus strain AD169 complete genome.	human herpesvirus 5	39,854	10-Feb-99
GB_HTG1:CEY39F1	178443	Z98861	178443	Caenorhabditis elegans chromosome II clone Y38F1, *** SEQUENCING IN PROGRESS ***. in unordered pieces.	Caenorhabditis elegans	35,880	03-DEC-1998

TABLE 4: ALIGNMENT RESULTS

GB_HTG1:CEY38F1	178443	Z98861	178443	GB_PR2:CNS00004	205573	AL049778	205573	GB_PR2:CNS00004	205573	AL049778	205573	Caenorhabditis elegans chromosome II clone Y38F1, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	35,880	03-DEC-1998
GB_PR3:HS934G17	107603	AL021155	107603	GB_GSS1:AG000894	723	AG000894	723	GB_GSS1:AG000894	723	AG000894	723	Homo sapiens DNA sequence from PAC 934G17 on chromosome 1p36.21. Contains the alternatively spliced CLCN6 gene for chloride channel proteins CLC-6A (KIAA0046) -B, -C and -D, the alternatively spliced NPPA gene coding for Atrial Natriuretic Factor ANF precursor (Atrial Natriuretic peptide ANP, Preproatrialatriatin), the NPPB gene for Brain Natriuretic Protein BNP, and a pseudogene similar to SBF1 (and other Myotubularin-related protein genes). Contains ESTs, STSs and the genomic marker D1S2740, complete sequence.	Homo sapiens	38,489	23-Nov-99
GB_EST26:AA899042	505	AA899042	505	GB_EST26:AA899042	505	AA899042	505	GB_EST26:AA899042	505	AA899042	505	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-643C12 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.	Homo sapiens	33,138	17-Jun-99
GB_RO:RNU09256	2098	U09256	2098	GB_GSS1:AG000894	723	AG000894	723	GB_GSS1:AG000894	723	AG000894	723	Trypanosoma simiae mini-exon DNA.	Homo sapiens	37,391	6-Feb-99
GB_EST29:AI599628	510	AI599628	510	GB_EST29:AI599628	510	AI599628	510	GB_EST29:AI599628	510	AI599628	510	UI-R-E0-bz-a-06-0-UI s2 UI-R-E0 Rattus norvegicus cDNA clone UI-R-E0-bz-a-06-0-UI 3' similar to g1485266(gb)U09256[RNU09256 Rattus norvegicus Sprague-Dawley transketolase mRNA, complete cds, mRNA sequence.	Trypanosoma simiae	35,135	15-Feb-99
GB_PL2:SPU66305	8226	U66305	8226	GB_PL2:SPU66305	8226	U66305	8226	GB_PL2:SPU66305	8226	U66305	8226	EST251331 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone REMEH65 3' end, mRNA sequence.	Rattus norvegicus	41,400	4-Jul-99
GB_PL1:SPBC2582	26016	AL031853	26016	GB_PL1:SPBC2582	26016	AL031853	26016	GB_PL1:SPBC2582	26016	AL031853	26016	Schizosaccharomyces pombe ABC transporter (mam1) gene, complete cds.	Rattus norvegicus	39,652	11-MAY-1994
GB_PL1:SPBC2G5	23645	AL033385	23645	GB_PL1:SPBC2G5	23645	AL033385	23645	GB_PL1:SPBC2G5	23645	AL033385	23645	S. pombe chromosome II cosmid c25B2.	Rattus sp.	39,096	21-Apr-99
GB_BA2:RRU65510	16259	U65510	16259	GB_BA2:RRU65510	16259	U65510	16259	GB_BA2:RRU65510	16259	U65510	16259	Schizosaccharomyces pombe	Schizosaccharomyces pombe	36,842	30-Jul-97
GB_PL1:LETHM27	1152	X95296	1152	GB_PL1:LETHM27	1152	X95296	1152	GB_PL1:LETHM27	1152	X95296	1152	S. pombe chromosome II cosmid c2G5.	Schizosaccharomyces pombe	36,803	09-OCT-1998
GB_EST38:AW033855	646	AW033855	646	GB_EST38:AW033855	646	AW033855	646	GB_EST38:AW033855	646	AW033855	646	Rhodospirillum rubrum CO-induced hydrogenase operon (cooM, cooK, cooL, cooX, cooU, cooH) genes, iron sulfur protein (cooF) gene, carbon monoxide dehydrogenase (cooS) gene, carbon monoxide dehydrogenase accessory proteins (cooC, cooT, cooJ) genes, putative transcriptional activator (cooA) gene, nicotinate-nucleotide pyrophosphorylase (nadC) gene, complete cds, L-aspartate oxidase (nadB) gene, and alkyl hydroperoxide reductase (ahpC) gene, partial cds.	Schizosaccharomyces pombe	36,803	04-DEC-1998
GB_EST34:AI785570	454	AI785570	454	GB_EST34:AI785570	454	AI785570	454	GB_EST34:AI785570	454	AI785570	454	L. esculentum mRNA for THM27 protein.	Rhodospirillum rubrum	41,970	9-Apr-97
GB_EST25:AI256147	684	AI256147	684	GB_EST25:AI256147	684	AI256147	684	GB_EST25:AI256147	684	AI256147	684	EST277426 tomato callus, TAMU Lycopersicon esculentum cDNA clone cLEC29F6 similar to transcription factor, myb-related, mRNA sequence.	Lycopersicon esculentum	38,919	10-Jun-96
GB_BA1:CARCG12	2079	X14979	2079	GB_BA1:CARCG12	2079	X14979	2079	GB_BA1:CARCG12	2079	X14979	2079	uj44d03.x1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1922789 3' similar to gb:Z28407.60S RIBOSOMAL PROTEIN L8 (HUMAN); mRNA sequence.	Lycopersicon esculentum	35,945	15-Sep-99
GB_EST15:AA497266	456	AA497266	456	GB_EST15:AA497266	456	AA497266	456	GB_EST15:AA497266	456	AA497266	456	u95e12.x1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1890190 3' similar to gb:Z28407.60S RIBOSOMAL PROTEIN L8 (HUMAN); mRNA sequence.	Mus musculus	37,565	2-Jul-99
GB_EST36:AI884217	515	AI884217	515	GB_EST36:AI884217	515	AI884217	515	GB_EST36:AI884217	515	AI884217	515	C. aurantiacus reaction center genes 1 and 2.	Mus musculus	41,232	12-Nov-98
GB_HTG1:CEY43F8_1	110000	Z95393	110000	GB_HTG1:CEY43F8_1	110000	Z95393	110000	GB_HTG1:CEY43F8_1	110000	Z95393	110000	fa04f08.s1 Zebrafish ICRFzfs Danio rerio cDNA clone 3A13 3', mRNA sequence.	Chloroflexus aurantiacus	36,943	23-Apr-91
GB_GSS8:AQ013755	715	AQ013755	715	GB_GSS8:AQ013755	715	AQ013755	715	GB_GSS8:AQ013755	715	AQ013755	715	fc75e10.x1 Zebrafish WashU MPIMG EST Danio rerio cDNA 3', mRNA sequence.	Danio rerio	38,631	30-Jun-97
GB_GSS3:B86449	434	B86449	434	GB_GSS3:B86449	434	B86449	434	GB_GSS3:B86449	434	B86449	434	Caenorhabditis elegans chromosome V clone Y43F8, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Danio rerio	38,012	26-Jul-99
GB_GSS5:AQ797072	449	AQ797072	449	GB_GSS5:AQ797072	449	AQ797072	449	GB_GSS5:AQ797072	449	AQ797072	449	RPCI11-23F24, TKBF RPCI-11 Homo sapiens genomic clone RPCI-11-23F24, genomic survey sequence.	Caenorhabditis elegans	37,889	Z95393
GB_BA2:AF011544	7527	AF011544	7527	GB_BA2:AF011544	7527	AF011544	7527	GB_BA2:AF011544	7527	AF011544	7527	RPCI11-23F24, TV RPCI-11 Homo sapiens genomic clone RPCI-11-23F24, genomic survey sequence.	Homo sapiens	41,724	14-Apr-99
GB_EST36:AI884217	515	AI884217	515	GB_EST36:AI884217	515	AI884217	515	GB_EST36:AI884217	515	AI884217	515	nbxb0071D10f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0071D10f, genomic survey sequence.	Homo sapiens	42,936	9-Apr-99
GB_EST36:AI884217	515	AI884217	515	GB_EST36:AI884217	515	AI884217	515	GB_EST36:AI884217	515	AI884217	515	nbxb0071D10f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0071D10f, genomic survey sequence.	Oryza sativa	39,101	4-Aug-99
GB_EST36:AI884217	515	AI884217	515	GB_EST36:AI884217	515	AI884217	515	GB_EST36:AI884217	515	AI884217	515	Bacillus subtilis phosphoribosylaminimidazole-carboxamide formyltransferase (purH-J) gene, partial cds, phosphoribosylglycinamide synthetase (purD), YecA (yecA), putative adenine deaminase (yecB), YecC (yecC), and YecD (yecD) genes, complete cds, and putative glutamate synthase (yecE) gene, partial cds.	Bacillus subtilis	36,927	06-OCT-1997

TABLE 4: ALIGNMENT RESULTS

	GB_BA2:AF011544	7527	AF011544	Bacillus subtilis phosphoribosylaminoimidazole-carboxamide formyltransferase (purH-J) gene, partial cds, phosphoribosylglycinamide synthetase (purD), YecA (yecA), putative adenine deaminase (yecB), YecC (yecC), and YecD (yecD) genes, complete cds, and putative glutamate synthase (yecE) gene, partial cds.	Bacillus subtilis	39,752	06-OCT-1997
rx00713	1056						
	GB_PAT:192037	241	I92037	Sequence 4 from patent US 5726299.	Unknown.	99,048	01-DEC-1998
	GB_PAT:178748	241	I78748	Sequence 4 from patent US 5693781.	Unknown.	99,048	3-Apr-98
	GB_HTG3:AC009281	221178	AC009281	Homo sapiens chromosome 15 clone 8_C_22 map 15, *** SEQUENCING IN PROGRESS *** , 49 unordered pieces.	Homo sapiens	36,255	12-Aug-99
rx00714	684						
	GB_PL1:CCR5839	871	AJ005839	Cyclotella cryptica mRNA for fucoxanthin chlorophyll a/c binding protein. fcp12.	Cyclotella cryptica	36,364	30-Jul-98
	GB_PR2:HS1002M8	111768	AL035454	Human DNA sequence from clone 1002M8 on chromosome 20p11.21-11.23, complete sequence.	Homo sapiens	36,444	23-Nov-99
	GB_PR2:HS1002M8	111768	AL035454	Human DNA sequence from clone 1002M8 on chromosome 20p11.21-11.23, complete sequence.	Homo sapiens	34,894	23-Nov-99
rx00716	636						
	GB_PAT:178753	1187	I78753	Sequence 9 from patent US 5693781.	Unknown.	36,022	3-Apr-98
	GB_PAT:192042	1187	I92042	Sequence 9 from patent US 5726299.	Unknown.	36,022	01-DEC-1998
	GB_HTG3:AC005769	200000	AC005769	Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS *** , 5 unordered pieces.	Homo sapiens	36,745	21-Aug-99
rx00719	1752						
	GB_BA2:U32687	11847	U32687	Haemophilus influenzae Rd section 2 of 163 of the complete genome.	Haemophilus influenzae Rd	36,937	29-MAY-1998
	GB_EST13:AA333602	357	AA333602	EST37710 Embryo, 8 week 1 Homo sapiens cDNA 5' end similar to guanine nucleotide-binding protein rap2, ras-oncogene related. mRNA sequence.	Homo sapiens	45,938	21-Apr-97
	GB_BA2:U32687	11847	U32687	Haemophilus influenzae Rd section 2 of 163 of the complete genome.	Haemophilus influenzae Rd	36,390	29-MAY-1998
rx00720	789						
	GB_EST1:M61974	437	M61974	EST00024 Fetal brain, Stratagene (cat#936206) Homo sapiens cDNA clone HFBA87. mRNA sequence.	Homo sapiens	40,138	26-MAY-1992
	GB_EST3:R73776	389	R73776	y15sh07.r1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:143197 5', mRNA sequence.	Homo sapiens	41,818	5-Jun-95
	GB_EST35:AL043192	793	AL043192	DKFZp434G0723_r1 434 (synonym: hles3) Homo sapiens cDNA clone DKFZp434G0723 5', mRNA sequence.	Homo sapiens	38,571	29-Sep-99
rx00722	1088						
	GB_HTG3:AC008573	205755	AC008573	Homo sapiens chromosome 5 clone CIT-HSPC_551111, *** SEQUENCING IN PROGRESS *** , 95 unordered pieces.	Homo sapiens	38,506	3-Aug-99
	GB_HTG3:AC008573	205755	AC008573	Homo sapiens chromosome 5 clone CIT-HSPC_551111, *** SEQUENCING IN PROGRESS *** , 95 unordered pieces.	Homo sapiens	38,506	3-Aug-99
	GB_BA1:MTV014	58280	AL021646	Mycobacterium tuberculosis H37Rv complete genome; segment 137/162.	Mycobacterium tuberculosis	41,392	18-Jun-98
rx00724	2100						
	GB_BA1:SC7A1	32039	AL034447	Streptomyces coelicolor cosmid 7A1.	Streptomyces coelicolor	54,858	15-DEC-1998
	GB_BA1:BSY13937	27779	Y13937	Bacillus subtilis genomic DNA from the spoVM region.	Bacillus subtilis	47,010	30-MAR-1998
	GB_BA2:L78127	1225	L78127	Enterococcus faecium genomic DNA fragment.	Enterococcus faecium	36,880	18-Aug-99
rx00726	614						
	GB_BA1:BACJH642	282700	D84432	Bacillus subtilis DNA, 283 Kb region containing skin element.	Bacillus subtilis	56,694	6-Feb-99
	GB_BA1:BSUB0013	218470	Z99116	Bacillus subtilis complete genome (section 13 of 21): from 2395261 to 2613730.	Bacillus subtilis	36,513	26-Nov-97
rx00729							
	GB_BA1:SC4H8	15560	AL020958	Streptomyces coelicolor cosmid 4H8.	Streptomyces coelicolor	35,073	10-DEC-1997
rx00730	930						
	GB_HTG3:AC010758	145821	AC010758	Homo sapiens clone 1_B_18, *** SEQUENCING IN PROGRESS *** , 20 unordered pieces.	Homo sapiens	35,738	22-Sep-99
	GB_HTG3:AC010758	145821	AC010758	Homo sapiens clone 1_B_18, *** SEQUENCING IN PROGRESS *** , 20 unordered pieces.	Homo sapiens	35,738	22-Sep-99
	GB_GSS13:AQ469090	414	AQ469090	CITBI-E1-2596D12.TF CITBI-E1 Homo sapiens genomic clone 2596D12, genomic survey sequence.	Homo sapiens	36,842	23-Apr-99
rx00731	2619						
	GB_BA1:CGLYSI	4232	X60312	C. glutamicum lysI gene for L-lysine permease.	Corynebacterium glutamicum	100,000	30-Jan-92

**TABLE 4: ALIGNMENT RESULTS**

GB_BA1:CGLYSI	4232	X60312	C. glutamicum lysl gene for L-lysine permease.	37,645	Corynebacterium glutamicum	30-Jan-92
GB_BA1:MTCY10G2	38970	Z92539	Mycobacterium tuberculosis H37Rv complete genome; segment 47/162.	54,427	Mycobacterium tuberculosis	17-Jun-98
GB_HTG1:HSJ564F22	106277	AL080249	Homo sapiens chromosome 20 clone RP4-564F22, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	44,000	Homo sapiens	23-Nov-99
GB_HTG1:HSJ564F22	106277	AL080249	Homo sapiens chromosome 20 clone RP4-564F22, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	44,000	Homo sapiens	23-Nov-99
GB_PL2:AF100167	1557	AF100167	Glycine max unknown mRNA.	35,823	Glycine max	4-Nov-98
GB_EST28:AI465702	268	AI465702	ww83g01.y1 Stralagene mouse skin (#937313) Mus musculus cDNA clone IMAGE:1261584 5' mRNA sequence.	43,226	Mus musculus	09-MAR-1999
GB_EST20:AA856157	359	AA856157	ww83g01.r1 Stralagene mouse skin (#937313) Mus musculus cDNA clone IMAGE:1261584 5' mRNA sequence.	43,226	Mus musculus	06-MAR-1998
GB_HTG2:AC007185	199340	AC007185	Drosophila melanogaster chromosome 2 clone BACR44N04 (D545) RPCI-98 44.N.4 map 36A-36A strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 50 unordered pieces.	39,583	Drosophila melanogaster	2-Aug-99
GB_HTG2:AC007185	199340	AC007185	Drosophila melanogaster chromosome 2 clone BACR44N04 (D545) RPCI-98 44.N.4 map 36A-36A strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 50 unordered pieces.	39,583	Drosophila melanogaster	2-Aug-99
GB_PL2:F17123	134784	AF160182	Arabidopsis thaliana BAC F17123.	37,788	Arabidopsis thaliana	20-Jun-99
GB_IN1:CEC05C10	26263	Z48178	Caenorhabditis elegans cosmid C05C10, complete sequence.	39,243	Caenorhabditis elegans	2-Sep-99
GB_IN1:CEC05C10	26263	Z48178	Caenorhabditis elegans cosmid C05C10, complete sequence.	38,041	Caenorhabditis elegans	2-Sep-99
GB_GSS8:AAQ093649	320	AQ093649	HS_3022_A1_E06_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3022 Col=11 Row=I, genomic survey sequence.	34,277	Homo sapiens	27-Aug-98
GB_GSS9:AAQ093649	320	AQ093649	HS_3022_A1_E06_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3022 Col=11 Row=I, genomic survey sequence.	34,277	Homo sapiens	27-Aug-98
GB_HTG2:AC008195	130309	AC008195	Drosophila melanogaster chromosome 3 clone BACR42120 (D748) RPCI-98 42.1.20 map 93F-93F strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 77 unordered pieces.	38,095	Drosophila melanogaster	2-Aug-99
GB_IN2:DMU53190	3477	U53190	Drosophila melanogaster Camguk (cmg) mRNA, complete cds.	39,623	Drosophila melanogaster	30-Nov-98
GB_HTG2:AC008195	130309	AC008195	Drosophila melanogaster chromosome 3 clone BACR42120 (D748) RPCI-98 42.1.20 map 93F-93F strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 77 unordered pieces.	38,095	Drosophila melanogaster	2-Aug-99
GB_HTG3:AC010076	148614	AC010076	Homo sapiens chromosome 15 clone BAC 64K10 map 14q25, LOW-PASS SEQUENCE SAMPLING.	36,336	Homo sapiens	11-Sep-99
GB_HTG3:AC010076	148614	AC010076	Homo sapiens chromosome 15 clone BAC 64K10 map 14q25, LOW-PASS SEQUENCE SAMPLING.	36,336	Homo sapiens	11-Sep-99
GB_PR3:HS402G11	177241	AL022328	Human DNA sequence from clone 402G11 on chromosome 22q13.31-13.33 Contains genes for SAPK3 (stress-activated protein kinase 3), PRKM11 (protein kinase mitogen-activated 11), KIAA0315, ESTs, GSSs and CpG islands, complete sequence.	38,752	Homo sapiens	23-Nov-99
GB_HTG4:AC010081	176777	AC010081	Homo sapiens clone NH0065E07, *** SEQUENCING IN PROGRESS ***; 1 unordered pieces.	37,016	Homo sapiens	29-OCT-1999
GB_HTG4:AC010081	176777	AC010081	Homo sapiens clone NH0065E07, *** SEQUENCING IN PROGRESS ***; 1 unordered pieces.	37,016	Homo sapiens	29-OCT-1999
GB_HTG3:AC011194	196098	AC011194	Mus musculus chromosome 11 clone 196_F_5 map 11, *** SEQUENCING IN PROGRESS ***; 32 unordered pieces.	38,735	Mus musculus	01-OCT-1999
GB_GSS13:AAQ457887	478	AQ457887	HS_5189_B2_B06_SPEE RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=765 Col=12 Row=D, genomic survey sequence.	37,844	Homo sapiens	23-Apr-99
GB_IN1:EHEXRIDRI	170	X58630	E.histolytica extrachromosomal ribosomal DNA for DRA I repeat unit.	42,353	Entamoeba histolytica	13-Aug-91
GB_IN1:EHEXRIDNA	3699	X61182	E.histolytica extrachromosomal ribosomal DNA downstream of rRNA genes.	38,905	Entamoeba histolytica	2-Sep-96
GB_BA1:BSUB0003	209100	Z99106	Bacillus subtilis complete genome (section 3 of 21); from 402751 to 611850.	37,238	Bacillus subtilis	26-Nov-97
GB_BA1:AB001488	148068	AB001488	Bacillus subtilis genome sequence, 148 kb sequencing of the region between 35 and 47 degree.	37,238	Bacillus subtilis	13-Feb-99
GB_HTG3:AC008060	161486	AC008060	Homo sapiens clone DJ0912113, *** SEQUENCING IN PROGRESS ***; 4 unordered pieces.	39,474	Homo sapiens	13-Aug-99

TABLE 4: ALIGNMENT RESULTS

rx00751	951	GB_PR4:AC006449	286758	AC006449	Homo sapiens chromosome 17, clone hCIT 58_E_17, complete sequence.	Homo sapiens	38,223	23-OCT-1999
		GB_HTG2:AC002118	170891	AC002118	Homo sapiens chromosome 17 clone 303_E_14, *** SEQUENCING IN PROGRESS ***; 20 unordered pieces.	Homo sapiens	37,112	13-Feb-98
		GB_HTG2:AC002118	170891	AC002118	Homo sapiens chromosome 17 clone 303_E_14, *** SEQUENCING IN PROGRESS ***; 20 unordered pieces.	Homo sapiens	37,112	13-Feb-98
rx00752	552	GB_PL2:F5K24	109786	AF128395	Arabidopsis thaliana BAC F5K24.	Arabidopsis thaliana	38,899	03-MAR-1999
		GB_BA1:MBHRDED	6300	Y09870	M.barkeri hdrE & hdrD genes, ORF1, ORF2, ORF3 & ORF4.	Methanosarcina barkeri	40,609	04-DEC-1998
		GB_PL1:SC9920	23498	Z48639	S.cerevisiae chromosome XIII cosmid 9920.	Saccharomyces cerevisiae	35,754	11-Aug-97
rx00757	1377	GB_PAT:E13655	2260	E13655	gDNA encoding glucose-6-phosphate dehydrogenase.	Corynebacterium glutamicum	46,045	24-Jun-98
		GB_GSS9:AQ103710	369	AQ103710	HS_3092_B1_C01_MF_CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3092 Col=1 Row=F, genomic survey sequence.	Homo sapiens	36,339	27-Aug-98
rx00763	906	GB_HTG3:AC009305	167705	AC009305	Homo sapiens clone NH0153B21, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Homo sapiens	36,691	13-Aug-99
		GB_BA1:SC787	13800	AL009199	Streptomyces coelicolor cosmid 787.	Streptomyces coelicolor	39,013	02-DEC-1997
		GB_HTG2:HSJ473J16	203460	AL109942	Homo sapiens chromosome 6 clone RP3-473J16 map q25.3-26, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,192	03-DEC-1999
		GB_HTG2:HSJ473J16	203460	AL109942	Homo sapiens chromosome 6 clone RP3-473J16 map q25.3-26, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,192	03-DEC-1999
rx00765	810	GB_BA1:MTV043	68848	AL022004	Mycobacterium tuberculosis H37Rv complete genome; segment 40/162.	Mycobacterium tuberculosis	38,568	24-Jun-99
		GB_BA2:PAU93274	8008	U93274	Pseudomonas aeruginosa YafE (yafE), LeuB (leuB), Asd (asd), FimV (fimV), and HisT (hisT) genes, complete cds; TrpF (trpF) gene, partial cds, and unknown gene.	Pseudomonas aeruginosa	37,656	23-Jun-98
		GB_BA1:MTCY31	37630	Z73101	Mycobacterium tuberculosis H37Rv complete genome; segment 41/162.	Mycobacterium tuberculosis	38,209	17-Jun-98
rx00768	1242	GB_HTG5:AC008194	194555	AC008194	Drosophila melanogaster chromosome X clone BACR49A05 (D745) RPCI-98.49.A.5 map 18A-18A strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 90 unordered pieces.	Drosophila melanogaster	34,078	15-Nov-99
		GB_HTG5:AC008194	194555	AC008194	Drosophila melanogaster chromosome X clone BACR49A05 (D745) RPCI-98.49.A.5 map 18A-18A strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 90 unordered pieces.	Drosophila melanogaster	31,194	15-Nov-99
		GB_BA2:AF044495	9599	AF044495	Agrobacterium tumefaciens chemotaxis operon, complete sequence.	Agrobacterium tumefaciens	40,165	2-Jul-98
rx00769	336	GB_PR3:AC003068	42184	AC003068	Human Cosmid g5129z059 from 7q31.3, complete sequence.	Homo sapiens	35,152	6-Nov-97
		GB_PR2:HSAC000374	41585	AC000374	Human cosmid g1980a170, complete sequence.	Homo sapiens	35,152	12-MAR-1997
		GB_PR3:AC003068	42184	AC003068	Human Cosmid g5129z059 from 7q31.3, complete sequence.	Homo sapiens	37,309	6-Nov-97
rx00771	942	GB_PR2:HS172K2	131234	Z84814	Human DNA sequence from PAC 172K2 on chromosome 6 contains HLA CLASS II DRA pseudogene, DRB3*01012 genes, DRB9 pseudogene butyrophilin precursor and ESTs.	Homo sapiens	34,719	23-Nov-99
		GB_HTG1:HSA555E18	1177	AL121780	Homo sapiens chromosome 20 clone RP11-555E18, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	41,450	23-Nov-99
		GB_HTG1:HSA555E18	1177	AL121780	Homo sapiens chromosome 20 clone RP11-555E18, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	41,450	23-Nov-99
rx00781	411	GB_PR3:HS48A11	129294	AL031132	Human DNA sequence from clone 48A11 on chromosome 20p12 Contains EST, STS, GSS, complete sequence.	Homo sapiens	37,284	23-Nov-99
		GB_IN1:CELC03B1	42297	U40952	Caenorhabditis elegans cosmid C03B1.	Caenorhabditis elegans	38,177	25-Nov-95
		GB_PR3:HS48A11	129294	AL031132	Human DNA sequence from clone 48A11 on chromosome 20p12 Contains EST, STS, GSS, complete sequence.	Homo sapiens	35,162	23-Nov-99
rx00785	680	GB_EST11:AA223451	349	AA223451	z106c001.r1 Striatagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone IMAGE:650689 5', mRNA sequence.	Homo sapiens	38,682	19-Feb-97
		GB_EST9:AA081255	446	AA081255	zn33d08.r1 Striatagene endothelial cell 937223 Homo sapiens cDNA clone IMAGE:549231 5', mRNA sequence.	Homo sapiens	40,271	21-OCT-1996

**TABLE 4: ALIGNMENT RESULTS**

rx00788	348	GB_EST9:C16722	314	C16722	C16722	C16722 Clontech human aorta polyA+ mRNA (#6572) Homo sapiens cDNA clone GEN-522C02 5' mRNA sequence.	Homo sapiens	44,013	30-Sep-96
		GB_PL2:UMU62738	13812	U62738	U62738	Ustilago maydis siderophore peptide synthetase (sid2) gene, complete cds.	Ustilago maydis	31,792	30-DEC-1997
		GB_PR1:AB012723	40850	AB012723	AB012723	Homo sapiens gene for kinesin-like protein, complete cds.	Homo sapiens	35,398	8-Jan-99
		GB_HTG3:AC008625	16830	AC008625	AC008625	Homo sapiens chromosome 5 clone C17978SKB_157D17, *** SEQUENCING IN PROGRESS ***, 19 unordered pieces.	Homo sapiens	42,560	3-Aug-99
rx00795	651	GB_IN2:AC003120	59991	AC003120	AC003120	Drosophila melanogaster DNA sequence (P1 D501523 (D34)), complete sequence.	Drosophila melanogaster	39,252	17-Jul-98
		GB_EST19:AA802212	574	AA802212	AA802212	GM04027 5prime GM Drosophila melanogaster ovary BlueScript Drosophila melanogaster cDNA clone GM04027 5prime, mRNA sequence.	Drosophila melanogaster	37,828	25-Nov-98
rx00804	567	GB_IN2:AF168467	4652	AF168467	AF168467	Drosophila melanogaster dual specificity kinase DYRK2 mRNA, complete cds.	Drosophila melanogaster	36,933	5-Aug-99
		GB_GSS12:AQ356039	499	AQ356039	AQ356039	CITBI-E1-2535P11.TR CITBI-E1 Homo sapiens genomic clone 2535P11, genomic survey sequence.	Homo sapiens	40,569	24-Jan-99
		GB_PR4:AC005037	190508	AC005037	AC005037	Homo sapiens clone NH0469M07, complete sequence.	Homo sapiens	41,209	14-MAY-1999
rx00805	1005	GB_HTG5:AC007272	175483	AC007272	AC007272	Homo sapiens clone NH0013J08, *** SEQUENCING IN PROGRESS ***, 5 unordered pieces.	Homo sapiens	41,209	2-Nov-99
		GB_GSS1:CNS00U61	320	AL090583	AL090583	Arabidopsis thaliana genome survey sequence SP6 end of BAC T6D17 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.	Arabidopsis thaliana	36,364	28-Jun-99
		GB_PL1:AB026639	63921	AB026639	AB026639	Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone: K21L13, complete sequence.	Arabidopsis thaliana	38,485	07-MAY-1999
		GB_PL1:AB026639	63921	AB026639	AB026639	Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone: K21L13, complete sequence.	Arabidopsis thaliana	35,451	07-MAY-1999
rx00808	1581	GB_BA1:MLCB2548	38916	AL023093	AL023093	Mycobacterium leprae cosmid B2548.	Mycobacterium leprae	50,854	27-Aug-99
		GB_BA1:MLCL373	37304	AL035500	AL035500	Mycobacterium leprae cosmid L373.	Mycobacterium leprae	40,295	27-Aug-99
		GB_PL2:SCE9781	68302	U18916	U18916	Saccharomyces cerevisiae chromosome V cosmid 9781, 8198, 9115, 9981, and lambda clones 3955 and 6052.	Saccharomyces cerevisiae	37,677	1-Aug-97
rx00812	1182	GB_HTG2:AC006003	114949	AC006003	AC006003	Homo sapiens clone DJ0782K24, *** SEQUENCING IN PROGRESS ***, 16 unordered pieces.	Homo sapiens	35,284	22-Nov-98
		GB_HTG2:AC006003	114949	AC006003	AC006003	Homo sapiens clone DJ0782K24, *** SEQUENCING IN PROGRESS ***, 16 unordered pieces.	Homo sapiens	35,284	22-Nov-98
		GB_GSS9:QA090529	323	AQ090529	AQ090529	HS_3007_B1_E09_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3007 Col=17 Row=J, genomic survey sequence.	Homo sapiens	41,176	26-Aug-98
rx00814	897	GB_VI:EHVU20824	184427	U20824	U20824	Equine herpesvirus 2, complete genome.	Equine herpesvirus 2	35,274	2-Feb-96
		GB_VI:EHVU20824	184427	U20824	U20824	Equine herpesvirus 2, complete genome.	Equine herpesvirus 2	38,808	2-Feb-96
		GB_PR3:HS466N1	79528	Z97630	Z97630	Human DNA sequence from clone 466N1 on chromosome 22q12-13 Contains H1F0(H1 histone family, member 0) gene, 2-amino-3-ketobutyrate-CoA ligase( nuclear gene encoding mitochondrial protein), GALR3 (galanin receptor) gene, ESTs, GSSs and CpG islands, complete sequence.	Homo sapiens	38,496	23-Nov-99
rx00815	696	GB_PR3:CNS01DRL	174928	AL117355	AL117355	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-354E14 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.	Homo sapiens	36,578	26-Nov-99
		GB_PR4:AC007283	127361	AC007283	AC007283	Homo sapiens clone NH0536118, complete sequence.	Homo sapiens	37,609	28-Sep-99
		GB_EST14:AA406984	477	AA406984	AA406984	MBAFCZ7H08T3 Brugia malayi adult female cDNA (SAW96MLW-BmAf) Brugia malayi cDNA clone AFCZ7H08 5' mRNA sequence.	Brugia malayi	41,919	01-MAY-1997
rx00816	420	GB_EST27:AL414036	467	AL414036	AL414036	ma03e08.x1 Soares mouse p3NMf19.5 Mus musculus cDNA clone IMAGE:303494 3' similar to TR:Q85299 Q85299 HOMOLOGUE OF RETROVIRAL PSEUDOPROTEASE. ; mRNA sequence.	Mus musculus	40,176	9-Feb-99
		GB_GSS15:AQ642295	501	AQ642295	AQ642295	RPCI93-DpnII-28G21 TV RPCI93-DpnII Trypanosoma brucei genomic clone RPCI93-DpnII-28G21, genomic survey sequence.	Trypanosoma brucei	37,540	8-Jul-99
		GB_PL2:ZMU82481	2750	U82481	U82481	Zea mays K1 domain interacting kinase 1 (KIK1) mRNA, complete cds.	Zea mays	41,783	1-Jan-98
rx00826	654	GB_PR4:AC008179	181745	AC008179	AC008179	Homo sapiens clone NH0576F01, complete sequence.	Homo sapiens	35,736	28-Sep-99
		GB_HTG1:AC002413	63369	AC002413	AC002413	Homo sapiens chromosome X clone bWXD111, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.	Homo sapiens	37,600	12-Aug-97
		GB_HTG1:AC002413	63369	AC002413	AC002413	Homo sapiens chromosome X clone bWXD111, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.	Homo sapiens	37,600	12-Aug-97
rx00830	846	GB_GSS6:AQ823465	535	AQ823465	AQ823465	HS_3217_A1_D08_TTC CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3217 Col=15 Row=G, genomic survey sequence.	Homo sapiens	40,417	26-Aug-99



## TABLE 4: ALIGNMENT RESULTS

rx00831	GB_GSS6:AQ825402	381	AQ825402	HS_5498_A1_G01_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1074 Col=1 Row=M, genomic survey sequence.	Homo sapiens	43,068	26-Aug-99
	GB_HTG1:HSU242F8	92944	AL022167	Homo sapiens chromosome X clone LLOXNC01-242F8, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	38,321	23-Nov-99
rx00835	GB_EST35:A1864917	468	A1864917	wj66f1.1.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2407821 3' similar to WP.F01F1.9 Homo sapiens CE01235 VACUOLAR AMINOPEPTIDASE 1, mRNA sequence.	Homo sapiens	37,607	30-Aug-99
	GB_EST35:A1864917	468	A1864917	wj66f1.1.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2407821 3' similar to WP.F01F1.9 Homo sapiens CE01235 VACUOLAR AMINOPEPTIDASE 1, mRNA sequence.	Homo sapiens	38,444	30-Aug-99
rx00836	GB_EST11:AA212728	424	AA212728	mw81g02.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:677138 5', mRNA sequence.	Mus musculus	40,284	18-Feb-97
	GB_EST26:A1390258	557	A1390258	mw81g02.y1 Soares mouse NML Mus musculus cDNA clone IMAGE:677138 5', mRNA sequence.	Mus musculus	41,261	2-Feb-99
rx00840	GB_PR3:AC003669	159446	AC003669	Homo sapiens Xp22 BAC GS-594A7 (Genome Systems Human BAC library) contains Bmx gene, complete sequence.	Homo sapiens	34,914	24-MAR-1998
rx00841	GB_BA1:U00017	42157	U00017	Mycobacterium leprae cosmid B2126.	Mycobacterium leprae	35,635	01-MAR-1994
	GB_BA1:MLCB2533	40245	AL035310	Mycobacterium leprae cosmid B2533.	Mycobacterium leprae	38,280	27-Aug-99
rx00853	GB_RO:AB022047S7	18721	AB022053	Mus musculus gene for prollyl oligopeptidase, exon 11, 12, 13, 14, 15 and complete cds.	Mus musculus	36,633	20-Aug-99
	GB_PR3:HS531H16	155116	AL031684	Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 531H16, complete sequence.	Homo sapiens	41,110	23-Nov-99
rx00854	GB_PR3:HS531H16	155116	AL031664	Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 531H16, complete sequence.	Homo sapiens	37,343	23-Nov-99
	GB_HTG3:AC010264	81671	AC010284	Homo sapiens chromosome 5 clone CIT-HSPC_468K18, *** SEQUENCING IN PROGRESS ***; 66 unordered pieces.	Homo sapiens	38,776	15-Sep-99
rx00855	GB_IN1:CELM04G7	41778	AF036700	Caenorhabditis elegans cosmid M04G7.	Caenorhabditis elegans	37,349	05-DEC-1997
	GB_EST20:AA850405	451	AA850405	EST193172 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone ROVAF27 3' end, mRNA sequence.	Rattus sp.	40,789	30-Apr-98
rx00855	GB_HTG2:AF165144	110891	AF165144	Homo sapiens chromosome 8 clone BAC 393A07 map 8q, *** SEQUENCING IN PROGRESS *** in ordered pieces.	Homo sapiens	34,234	16-Jul-99
	GB_HTG2:AC007173	140775	AC007173	Drosophila melanogaster chromosome 2 clone BACR01A03 (D538) RPCI-98 01.A.3 map 36E-36E strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 36 unordered pieces.	Drosophila melanogaster	36,341	2-Aug-99
rx00855	GB_HTG2:AC007173	140775	AC007173	Drosophila melanogaster chromosome 2 clone BACR01A03 (D538) RPCI-98 01.A.3 map 36E-36E strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 36 unordered pieces.	Drosophila melanogaster	36,341	2-Aug-99
	GB_PL2:YSC8179	44113	U00062	Saccharomyces cerevisiae chromosome VIII cosmid 8179.	Saccharomyces cerevisiae	38,560	4-Sep-97
rx00861	GB_BA1:CGORF4GEN	2398	X95649	C. glutamicum ORF4 gene.	Corynebacterium glutamicum	100,000	10-MAR-1998
	GB_BA1:SC9A10	9000	AL031260	Streptomyces coelicolor cosmid 9A10.	Streptomyces coelicolor	63,830	11-Aug-98
rx00862	GB_BA2:AF039028	2475	AF039028	Streptomyces toyocaensis D-ala-D-ala dipeptidase (varXst) gene, complete cds; and unknown gene.	Streptomyces toyocaensis	61,939	5-Jan-99
	GB_PAT:E14520	2001	E14520	DNA encoding Brevibacterium dihydrodipicolinic acid synthase.	Corynebacterium glutamicum	36,154	28-Jul-99

TABLE 4: ALIGNMENT RESULTS

rx00869	1044	GB_PAT:12773	2001	E12773	DNA encoding Brevibacterium dihydrodipicolinic acid reductase.	Corynebacterium glutamicum	36,154	24-Jun-98
		GB_PAT:16749	2001	E16749	gDNA encoding dihydrodipicolinate synthase (DDPS).	Corynebacterium glutamicum	36,154	28-Jul-99
		GB_EST24:AI166579	645	AI166579	xylem.est.398 Poplar xylem Lambda ZAPII library Populus balsamifera subsp. trichocarpa cDNA 5' mRNA sequence.	Populus balsamifera subsp. trichocarpa	39,854	03-DEC-1998
		GB_BA1:MTCY06H11	38000	Z85982	Mycobacterium tuberculosis H37Rv complete genome; segment 73/162.	Mycobacterium tuberculosis	42,801	17-Jun-98
		GB_EST34:AV153098	283	AV153098	AV153098 Mus musculus hippocampus C57BL/6J adult Mus musculus cDNA clone 2900052L10, mRNA sequence.	Mus musculus	39,576	7-Jul-99
rx00874	1212	GB_HTG2:AC007885	108561	AC007885	Drosophila melanogaster chromosome 2 clone BACR02G15 (D643) RPCI-98 02.G.15 map 60F-60F strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 65 unordered pieces.	Drosophila melanogaster	38,276	2-Aug-99
		GB_HTG2:AC007582	127205	AC007582	Drosophila melanogaster chromosome 2 clone BACR17E16 (D642) RPCI-98 17.E.16 map 60F-60F strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 81 unordered pieces.	Drosophila melanogaster	36,246	2-Aug-99
		GB_HTG2:AC007885	108561	AC007885	Drosophila melanogaster chromosome 2 clone BACR02G15 (D643) RPCI-98 02.G.15 map 60F-60F strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 65 unordered pieces.	Drosophila melanogaster	38,276	2-Aug-99
rx00876	1878	GB_EST10:AA144736	479	AA144736	m72d08.r1 Stratiogene mouse testis (#937308) Mus musculus cDNA clone IMAGE:602991 5', mRNA sequence.	Mus musculus	41,474	11-Feb-97
		GB_EST32:AU069076	316	AU069076	AU069076 Rice callus Oryza sativa cDNA clone C51993_1A, mRNA sequence.	Oryza sativa	46,330	7-Jun-99
		GB_EST10:AA144736	479	AA144736	m72d08.r1 Stratiogene mouse testis (#937308) Mus musculus cDNA clone IMAGE:602991 5', mRNA sequence.	Mus musculus	43,243	11-Feb-97
rx00881	501	GB_HTG4:AC010103	192320	AC010103	Homo sapiens chromosome unknown clone NH0508H21, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	36,620	29-OCT-1999
		GB_HTG4:AC010103	192320	AC010103	Homo sapiens chromosome unknown clone NH0508H21, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	36,620	29-OCT-1999
		GB_HTG4:AC010103	192320	AC010103	Homo sapiens chromosome unknown clone NH0508H21, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	34,280	29-OCT-1999
rx00882	801	GB_BA1:MTCY48	35377	Z74020	Mycobacterium tuberculosis H37Rv complete genome; segment 69/162.	Mycobacterium tuberculosis	37,927	17-Jun-98
		GB_PAT:AR005211	3453	AR005211	Sequence 1 from patent US 5747651.	Unknown.	39,620	04-DEC-1998
		GB_PAT:140600	3453	140600	Sequence 1 from patent US 5621090.	Unknown.	39,620	13-MAY-1997
rx00883	642	GB_PR2:HS217016	87552	AL031771	Human DNA sequence from clone 217016 on chromosome 6q24 Contains GSS, complete sequence.	Homo sapiens	33,866	23-Nov-99
		GB_PR2:HS217016	87552	AL031771	Human DNA sequence from clone 217016 on chromosome 6q24 Contains GSS, complete sequence.	Homo sapiens	35,479	23-Nov-99
rx00887								
rx00889	711	GB_BA1:MTCY27	27548	Z95208	Mycobacterium tuberculosis H37Rv complete genome; segment 104/162.	Mycobacterium tuberculosis	36,978	17-Jun-98
		GB_BA1:U00016	42931	U00016	Mycobacterium leprae cosmid B1937.	Mycobacterium leprae	55,651	01-MAR-1994
		GB_PR4:AC007326	102898	AC007326	Homo sapiens, complete sequence.	Homo sapiens	40,205	2-Nov-99
rx00893	720	GB_PL1:HVCPIII	6225	Y09602	H. vulgare gene encoding serine carboxypeptidase II, CP-MII.	Hordeum vulgare	35,704	10-MAR-1997
		GB_EST35:AI814621	441	AI814621	wf75d04.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2408647 3' similar to TR:O00578 O00578 KIAA0167. [1], mRNA sequence.	Homo sapiens	37,788	24-Aug-99
		GB_EST3:R51723	376	R51723	yg77h06.r1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:39671 5' similar to gb:M77016 TROPOMODULIN (HUMAN); mRNA sequence.	Homo sapiens	41,489	18-MAY-1995
rx00895	714	GB_HTG3:AC009414	188673	AC009414	Homo sapiens clone NH0490M08, *** SEQUENCING IN PROGRESS ***; 5 unordered pieces.	Homo sapiens	36,775	17-Sep-99
		GB_HTG3:AC009414	188673	AC009414	Homo sapiens clone NH0490M08, *** SEQUENCING IN PROGRESS ***; 5 unordered pieces.	Homo sapiens	36,775	17-Sep-99
		GB_PR3:HSJ824F16	139330	AL050325	Human DNA sequence from clone 824F16 on chromosome 20, complete sequence.	Homo sapiens	37,286	23-Nov-99

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

rx00904	815	GB_HTG5:AC006447	141862	AC006447	Mus musculus, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Mus musculus	35,945	17-Nov-99
		GB_HTG5:AC011064	233428	AC011064	Drosophila melanogaster chromosome X clone BACN05G06 (D1107) RPCI-98 05.G.6 map 12F-13A strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 220 unordered pieces.	Drosophila melanogaster	37,783	16-Nov-99
		GB_HTG6:AC008334	154566	AC008334	Drosophila melanogaster chromosome X clone BACR08K05 (D885) RPCI-98 08.K.5 map 12F-12F strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 84 unordered pieces.	Drosophila melanogaster	37,783	02-DEC-1999
rx00908	681	GB_GSS12:AQ409791	561	AQ409791	HS_5090_B2_B12_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=666 Col=24 Row=D, genomic survey sequence.	Homo sapiens	39,711	17-MAR-1999
		GB_GSS3:B83773	535	B83773	CpG0110A CplOWagDNA1 Cryptosporidium parvum genomic, genomic survey sequence.	Cryptosporidium parvum	44,615	06-MAY-1999
		GB_GSS12:AQ409791	561	AQ409791	HS_5090_B2_B12_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=666 Col=24 Row=D, genomic survey sequence.	Homo sapiens	41,333	17-MAR-1999
rx00915	753	GB_HTG2:HS1118M15	190466	AL109964	Homo sapiens chromosome 20 clone RP5-1118M15, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	40,027	30-Nov-99
		GB_HTG2:HS1057B20	204291	AL109823	Homo sapiens chromosome 20 clone RP5-1057B20 map q11.21-12, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,535	30-Nov-99
		GB_HTG2:HS1118M15	190466	AL109964	Homo sapiens chromosome 20 clone RP5-1118M15, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	40,027	30-Nov-99
rx00916	3714	GB_EST28:AI543268	568	AI543268	SD09973.5prime SD Drosophila melanogaster Schneider L2 cell culture pOT2 Drosophila melanogaster cDNA clone SD09973 5prime, mRNA sequence.	Drosophila melanogaster	40,426	22-MAR-1999
		GB_IN2:AC004301	68620	AC004301	Drosophila melanogaster DNA sequence (P1 DS07134 (D192)), complete sequence	Drosophila melanogaster	37,696	29-MAY-1998
		GB_EST37:AI994315	524	AI994315	701502677 A. thaliana, Ohio State clone set Arabidopsis thaliana cDNA clone 701502677, mRNA sequence.	Arabidopsis thaliana	40,076	8-Sep-99
rx00917	2802	GB_BA1:SYCSLRB	146271	D64000	Synechocystis sp. PCC6803 complete genome, 19/27, 2392729-2538999.	Synechocystis sp.	38,447	13-Feb-99
		GB_HTG1:CEY39E4_2	110000	Z94158	Caenorhabditis elegans chromosome III clone Y39E4, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	38,218	Z94158
rx00921		GB_HTG1:CEY39E4_2	110000	Z94158	Caenorhabditis elegans chromosome III clone Y39E4, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	38,218	Z94158
rx00926	486	GB_OM:SSUJ75316	5996	U75316	Sus scrofa beta-myosin heavy chain mRNA, complete cds.	Sus scrofa	38,958	03-DEC-1996
		GB_EST21:AA970971	371	AA970971	op10b11.s1 NCI_CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1575261 3', mRNA sequence.	Homo sapiens	40,841	13-Apr-99
		GB_OM:SSUJ75316	5996	U75316	Sus scrofa beta-myosin heavy chain mRNA, complete cds.	Sus scrofa	38,578	03-DEC-1996
rx00930	876	GB_BA1:MTIC270A	1670	Z98045	Mycobacterium tuberculosis H37Rv complete genome; segment 116/162.	Mycobacterium tuberculosis	37,927	17-Jun-98
		GB_BA1:U00011	40429	U00011	Mycobacterium leprae cosmid B1177.	Mycobacterium leprae	38,623	01-MAR-1994
		GB_RO:S58745	817	S58745	thyrotroph embryonic factor=leucine zipper transcription factor [rats, pituitary, mRNA, 817 nt].	Rattus sp.	41,483	07-MAY-1993
rx00932	597	GB_PR4:AC009509	192690	AC009509	Homo sapiens 12p11-37.2-54.4 BAC RP11-1060J15 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	Homo sapiens	38,776	01-DEC-1999
		GB_PR3:AC004072	170658	AC004072	Human Chromosome X clone bWXD342, complete sequence.	Homo sapiens	35,000	08-MAR-1998
		GB_PR4:AC004617	176552	AC004617	Homo sapiens chromosome Y, clone 264.M.20, complete sequence.	Homo sapiens	35,702	13-OCT-1999
rx00933	585	GB_PL1:MGNGAGPOLH638	L35053	L35053	Transposon MAGGY gag and pol gene homologues, partial cds.	Magnaporthe grisea	40,283	4-Aug-94
		GB_PL1:MGNGAGPOLH638	L35053	L35053	Transposon MAGGY gag and pol gene homologues, partial cds.	Magnaporthe grisea	37,739	4-Aug-94
rx00940	519	GB_PR2:HS179N16	172048	Z95152	Homo sapiens DNA sequence from PAC 179N16 on chromosome 6p21.1-21.33. Contains the SAPK4 (MAPK p38delta) gene, and the alternatively spliced SAPK2 gene coding for CSaids binding protein CSBP2 and a MAPK p38beta LIKE protein. Contains ESTs, STSs and two predicted CpG islands, complete sequence.	Homo sapiens	38,252	23-Nov-99
		GB_EST26:AU001018	304	AU001018	Bombax mori p50(Daizo) Bombax mori cDNA clone fb09332f, mRNA sequence.	Bombax mori	45,745	15-Jan-99
		GB_EST26:AU001019	304	AU001019	AU001019 Bombax mori p50(Daizo) Bombax mori cDNA clone fb0934f, mRNA sequence.	Bombax mori	45,745	15-Jan-99
rx00943	1035	GB_BA2:AF079317	184457	AF079317	Sphingomonas aromaticivorans plasmid pNL1, complete plasmid sequence.	Sphingomonas aromaticivorans	38,151	12-Jan-99

**TABLE 4: ALIGNMENT RESULTS**

rx00946	897	GB_HTG3:AC008329	114408	AC008329	Drosophila melanogaster chromosome 2 clone BACR31D05 (D861) RPCI-98 31.D.5 map 28C-28D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 105 unordered pieces.	Drosophila melanogaster	34,317	17-Aug-99
		GB_HTG3:AC008329	114408	AC008329	Drosophila melanogaster chromosome 2 clone BACR31D05 (D861) RPCI-98 31.D.5 map 28C-28D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 105 unordered pieces.	Drosophila melanogaster	34,317	17-Aug-99
		GB_BA1:MTV008	63033	AL021246	Mycobacterium tuberculosis H37Rv complete genome; segment 108/162.	Mycobacterium tuberculosis	36,045	17-Jun-98
		GB_GSS14:AQ571765	526	AQ571765	HS_2094_A2_B09_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2094 Col=18 Row=C; genomic survey sequence.	Homo sapiens	38,021	1-Jun-99
		GB_RO:RRFE6G	2464	X60468	R. rattus FE65 gene for adaptor protein interacting with the beta-amyloid precursor protein intracellular domain.	Rattus rattus	38,417	1-Feb-96
rx00949	771	GB_VI:PPCCGAAA	5366	M26281	Hamster papovavirus complete genome.	Hamster papovavirus	36,579	22-MAY-1995
		GB_VI:HAPVXX	5366	X02449	Hamster Papovavirus (HapV) genome.	Hamster papovavirus	36,579	22-OCT-1999
		GB_BA2:AE000878	15432	AE000878	Methanobacterium thermoautotrophicum from bases 976801 to 992232 (section 84 of 148) of the complete genome.	Methanobacterium thermoautotrophicum	36,856	15-Nov-97
rx00959	579	GB_BA1:CGMTRAR	951	X75083	C. glutamicum mtrA gene locus with 5-methyltryptophan resistance.	Corynebacterium glutamicum	99,133	18-Aug-94
		GB_BA1:CGMTRA	587	X75084	C. glutamicum sequence corresponding to mtrA locus.	Corynebacterium glutamicum	99,216	18-Aug-94
		GB_BA1:BLTRP	7725	X04960	Brevibacterium lactofermentum tryptophan operon.	Corynebacterium glutamicum	96,800	10-Feb-99
rx00963	960	GB_EST15:AA484511	504	AA484511	n0807.s1 NCL_CGAP_Lit Homo sapiens cDNA clone IMAGE:913189 similar to gb:Y00764 UBQUINOL-CYTOCHROME C REDUCTASE 11 KD PROTEIN (HUMAN); mRNA sequence.	Homo sapiens	43,750	18-Aug-97
		GB_EST20:AA894481	544	AA894481	nw76b10.s1 NCL_CGAP_P12 Homo sapiens cDNA clone IMAGE:1252507 similar to gb:Y00764 UBQUINOL-CYTOCHROME C REDUCTASE 11 KD PROTEIN (HUMAN); mRNA sequence.	Homo sapiens	37,500	6-Apr-98
		GB_EST15:AA528497	582	AA528497	n196d07.s1 NCL_CGAP_P12 Homo sapiens cDNA clone IMAGE:984685 3' similar to gb:Y00764 UBQUINOL-CYTOCHROME C REDUCTASE 11 KD PROTEIN (HUMAN); mRNA sequence.	Homo sapiens	38,554	5-Aug-97
rx00969	1458	GB_BA1:CGHOMTHR	3685	Y00546	Corynebacterium glutamicum hom-thrB genes for homoserine dehydrogenase and homoserine kinase.	Corynebacterium glutamicum	99,588	12-Sep-93
		GB_PAT:109077	3685	I09077	Sequence 1 from Patent WO 8809819.	Unknown.	99,246	02-DEC-1994
		GB_BA1:BLTHRA	1483	Y00476	B. lactofermentum thr A gene.	Corynebacterium glutamicum	99,378	05-MAY-1993
rx00971	341	GB_BA1:CGHOMTHR	3685	Y00546	Corynebacterium glutamicum hom-thrB genes for homoserine dehydrogenase and homoserine kinase.	Corynebacterium glutamicum	35,435	12-Sep-93
		GB_PAT:109077	3685	I09077	Sequence 1 from Patent WO 8809819.	Unknown.	35,435	02-DEC-1994
		GB_BA1:BLTHRB	1139	Y00140	Brevibacterium lactofermentum thrB gene for homoserine kinase.	Corynebacterium glutamicum	40,964	12-Sep-93
rx00973	726	GB_BA1:CGHOMTHR	3685	Y00546	Corynebacterium glutamicum hom-thrB genes for homoserine dehydrogenase and homoserine kinase.	Corynebacterium glutamicum	41,797	12-Sep-93
		GB_PAT:109077	3685	I09077	Sequence 1 from Patent WO 8809819.	Unknown.	41,797	02-DEC-1994
		GB_IN2:AC006574	127035	AC006574	Drosophila melanogaster, chromosome 2R, region 39A3-39B1, P1 clones DS02919 and DS05130, complete sequence.	Drosophila melanogaster	37,355	16-Feb-99
rx00978	738	GB_PR2:HSAAC000372	41730	AC000372	Human cosmid g1980a186, complete sequence.	Homo sapiens	34,674	12-MAR-1997
		GB_PR3:AC005503	40998	AC005503	Homo sapiens clone UWGC:g5129s003 from 7q31, complete sequence.	Homo sapiens	34,674	20-Aug-98
		GB_PR2:HSAAC000372	41730	AC000372	Human cosmid g1980a186, complete sequence.	Homo sapiens	38,881	12-MAR-1997
rx00986	465	GB_GSS10:AQ258013	761	AQ258013	nbxb0019H05f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0019H05f, genomic survey sequence.	Oryza sativa	31,533	23-OCT-1998
		GB_PR3:HS83L6	61187	Z99130	Human DNA sequence from PAC 83L6 on chromosome Xp11.1-11.22. Contains ZXDA (ZFPA) zinc finger gene. ESTs and STSs, complete sequence.	Homo sapiens	38,395	23-Nov-99
		GB_PR3:HS58A24	96558	AL031115	Human DNA sequence from clone 58A24 on chromosome Xp11.1-11.23. Contains zinc finger X-linked proteins ZXDA, ZXDB, ESTs and STS, complete sequence.	Homo sapiens	37,333	23-Nov-99

TABLE 4: ALIGNMENT RESULTS

rx00987	588	GB_HTG1:HS24A17	2000	AL035452	Homo sapiens chromosome X clone RP6-24A17, *** SEQUENCING IN PROGRESS *** , in unordered pieces.	Homo sapiens	38,821	23-Nov-99
		GB_HTG1:HS24A17	2000	AL035452	Homo sapiens chromosome X clone RP6-24A17, *** SEQUENCING IN PROGRESS *** , in unordered pieces.	Homo sapiens	38,821	23-Nov-99
		GB_PR2:HS1156N12	146360	AL009047	Human DNA sequence from clone 1156N12 on chromosome X. Contains an STS and GSSs, complete sequence.	Homo sapiens	38,821	23-Nov-99
rx00988	546	GB_IN1:CELZC328	30350	AF000194	Caenorhabditis elegans cosmid ZC328.	Caenorhabditis elegans	36,122	23-Apr-97
		GB_IN1:CELZC328	30350	AF000194	Caenorhabditis elegans cosmid ZC328.	Caenorhabditis elegans	37,959	23-Apr-97
rx01005	969	GB_BA1:FVBPOAD2A	45519	D26094	Flavobacterium sp. plasmid pOAD2 DNA, whole sequence.	Flavobacterium sp.	37,998	6-Feb-99
		GB_GSS1:CNS00UGV	472	AL090973	Arabidopsis thaliana genome survey sequence SP6 end of BAC T8P9 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.	Arabidopsis thaliana	39,024	28-Jun-99
rx01007		GB_GSS1:CNS00S69	512	AL087999	Arabidopsis thaliana genome survey sequence SP6 end of BAC T1C4 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.	Arabidopsis thaliana	35,938	28-Jun-99
rx01008								
rx01011	1356	GB_EST38:AW039107	598	AW039107	EST281080 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA clone cLET12F19, mRNA sequence.	Lycopersicon esculentum	39,724	18-OCT-1999
		GB_BA1:MTY13E12	43401	Z95390	Mycobacterium tuberculosis H37Rv complete genome, segment 147/162.	Mycobacterium tuberculosis	38,618	17-Jun-98
		GB_BA1:MBU15140	2136	U15140	Mycobacterium bovis ribosomal proteins IF-1 (infA), L36 (rpmJ), S13 (rpsM) and S11 (rpsK) genes, complete cds, and S4 (rpsD) gene, partial cds.	Mycobacterium bovis	37,070	28-OCT-1996
rx01016	771	GB_BA1:CGBPHI16	962	Y12472	C. glutamicum DNA, attachment site bacteriophage Phi-16.	Corynebacterium glutamicum	45,098	05-MAR-1999
		GB_BA1:CGBPHI16	962	Y12472	C. glutamicum DNA, attachment site bacteriophage Phi-16.	Corynebacterium glutamicum	37,251	05-MAR-1999
rx01017	732	GB_BA1:CGBPHI16	962	Y12472	C. glutamicum DNA, attachment site bacteriophage Phi-16.	Corynebacterium glutamicum	39,245	05-MAR-1999
		GB_BA2:AF099014	2500	AF099014	Streptomyces coelicolor strain A3(2) transposase (tnpA) and Fe-containing superoxide dismutase I (sodF1) genes, complete cds.	Streptomyces coelicolor	38,036	1-Jun-99
		GB_HTG3:AC009249	119461	AC009249	Drosophila melanogaster chromosome 3 clone BACR02M06 (D1003) RPC1-98 02 M 6 map 98B-98B strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** , 97 unordered pieces.	Drosophila melanogaster	37,853	27-Aug-99
rx01021	622	GB_BA2:U39718	8603	U39718	Mycoplasma genitalium section 40 of 51 of the complete genome.	Mycoplasma genitalium	39,348	5-Nov-98
		GB_GSS3:BA6221	457	BA6221	HS-1063-A2-D06-MR-abi CIT Human Genomic Sperm Library C Homo sapiens genomic clone Plate=CT 796 Col=12 Row=G, genomic survey sequence.	Homo sapiens	39,933	21-OCT-1997
rx01023	1101	GB_OV:AF035529	848	AF035529	Xenopus laevis Smad6 mRNA, partial cds.	Xenopus laevis	37,203	1-Jan-98
		GB_HTG2:HSJ435K13	151301	AL109941	Homo sapiens chromosome 6 clone RP3-435K13 map q14.1-16.1, *** SEQUENCING IN PROGRESS *** , in unordered pieces.	Homo sapiens	34,405	03-DEC-1999
		GB_HTG2:HSJ435K13	151301	AL109941	Homo sapiens chromosome 6 clone RP3-435K13 map q14.1-16.1, *** SEQUENCING IN PROGRESS *** , in unordered pieces.	Homo sapiens	34,405	03-DEC-1999
rx01028	2172	GB_BA1:RCU57682	86896	U57682	Rhodobacter capsulatus cosmids 143-147, complete sequence.	Rhodobacter capsulatus	39,022	7-Feb-97
		GB_IN1:CBUS55260	2518	U55260	Caenorhabditis briggsae beta tubulin (mec-7) gene, complete cds.	Caenorhabditis briggsae	39,467	5-Jun-96
		GB_HTG1:CEY1A5	196643	AL008872	Caenorhabditis elegans chromosome III clone Y1A5, *** SEQUENCING IN PROGRESS *** , in unordered pieces.	Caenorhabditis elegans	38,168	9-Nov-97
		GB_HTG1:CEY1A5	196643	AL008872	Caenorhabditis elegans chromosome III clone Y1A5, *** SEQUENCING IN PROGRESS *** , in unordered pieces.	Caenorhabditis elegans	38,168	9-Nov-97

TABLE 4: ALIGNMENT RESULTS

rx01029	612	GB_PR3:HS466P17	149963	AL023806	Human DNA sequence from clone 466P17 on chromosome 6q24. Contains a putative novel gene, the 5' part of the EPM2A (Laforin) gene, ESTs, STSs, GSSs, genomic marker D6S1703 and D6S1443, a putative CpG island and a ca repeat polymorphism, complete sequence.	Homo sapiens	38,330	23-Nov-99
		GB_PR3:HS466P17	149963	AL023806	Human DNA sequence from clone 466P17 on chromosome 6q24. Contains a putative novel gene, the 5' part of the EPM2A (Laforin) gene, ESTs, STSs, GSSs, genomic marker D6S1703 and D6S1443, a putative CpG island and a ca repeat polymorphism, complete sequence.	Homo sapiens	39,262	23-Nov-99
rx01031	789	GB_RO:D78344	59641	D78344	Mouse DNA for Ig gamma-chains, partial cds.	Mus musculus	35,472	5-Feb-99
		GB_PR4:AC006948	168558	AC006948	Homo sapiens chromosome 17, clone hRPK_334_M_10, complete sequence.	Homo sapiens	44,005	27-Apr-99
		GB_PL2:AC011665	101845	AC011665	Arabidopsis thaliana chromosome 1 BAC T6L1 genomic sequence, complete sequence.	Arabidopsis thaliana	38,170	11-Nov-99
		GB_RO:MMU19724	5523	U19724	Mus musculus MMTV integration locus, aromatase gene, 3'UTR.	Mus musculus	35,256	17-Feb-96
rx01032	498	GB_EST9:AA118349	576	AA118349	mi56b06.r1 Stratagene mouse testis (#937308) Mus musculus cDNA clone IMAGE:515987 5' similar to gb:L04852 Mouse (MOUSE);, mRNA sequence.	Mus musculus	43,056	19-Nov-96
		GB_EST9:AA118349	576	AA118349	mi56b06.r1 Stratagene mouse testis (#937308) Mus musculus cDNA clone IMAGE:515987 5' similar to gb:L04852 Mouse (MOUSE);, mRNA sequence.	Mus musculus	42,273	19-Nov-96
rx01033	459	GB_GSS13:AQ434868	520	AQ434868	HS_5117_B1_D07_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=693 Col=13 Row=H, genomic survey sequence.	Homo sapiens	38,608	31-MAR-1999
		GB_HTG2:HSDJ79416	137124	AL108976	Homo sapiens chromosome 20 clone RP4-79416, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	28,929	27-Nov-99
		GB_HTG2:HSDJ79416	137124	AL108976	Homo sapiens chromosome 20 clone RP4-79416, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	28,929	27-Nov-99
rx01034	477	GB_PL2:ATT29H11	87011	AL049659	Arabidopsis thaliana DNA chromosome 3, BAC clone T29H11.	Arabidopsis thaliana	32,495	9-Jun-99
		GB_PL2:ATT29H11	87011	AL049659	Arabidopsis thaliana DNA chromosome 3, BAC clone T29H11.	Arabidopsis thaliana	40,042	9-Jun-99
		GB_EST25:AU045739	436	AU045739	AU045739 Mouse sixteen-cell-embryo cDNA Mus musculus cDNA clone J0940F02 3', mRNA sequence.	Mus musculus	35,435	09-DEC-1998
rx01035	729	GB_GSS1:CNS00QD8	526	AL085658	Arabidopsis thaliana genome survey sequence SP6 end of BAC F11C22 of IGF library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.	Arabidopsis thaliana	36,466	28-Jun-99
		GB_GSS13:AQ447948	515	AQ447948	mgxb0015A01r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0015A01r, genomic survey sequence.	Magnaporthe grisea	45,833	8-Apr-99
		GB_GSS1:CNS00QD8	526	AL085658	Arabidopsis thaliana genome survey sequence SP6 end of BAC F11C22 of IGF library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.	Arabidopsis thaliana	37,500	28-Jun-99
rx01036	576	GB_HTG2:AC004846	143577	AC004846	Homo sapiens clone DJ0647C14, *** SEQUENCING IN PROGRESS *** 21 unordered pieces.	Homo sapiens	38,137	12-Jun-98
		GB_EST19:AA804532	427	AA804532	ns28c05.s1 NCL_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1184936 3' similar to contains element MER40 repetitive element ;, mRNA sequence.	Homo sapiens	33,582	18-Feb-98
		GB_HTG2:AC006342	201618	AC006342	Homo sapiens clone DJ0054D12, *** SEQUENCING IN PROGRESS *** 3 unordered pieces.	Homo sapiens	38,137	11-Jan-99
rx01037	651	GB_PR4:AC004812	138532	AC004812	Homo sapiens PAC clone 267D11 from 12, complete sequence.	Homo sapiens	39,750	05-DEC-1998
		GB_HTG2:AC006342	201618	AC006342	Homo sapiens clone DJ0054D12, *** SEQUENCING IN PROGRESS *** 3 unordered pieces.	Homo sapiens	41,214	11-Jan-99
		GB_HTG2:AC004846	143577	AC004846	Homo sapiens clone DJ0647C14, *** SEQUENCING IN PROGRESS *** 21 unordered pieces.	Homo sapiens	41,214	12-Jun-98
rx01038								
rx01039	699	GB_PR4:HUAC004682	189134	AC004682	Homo sapiens Chromosome 16 BAC clone CIT987SK-A-259H10, complete sequence.	Homo sapiens	36,192	23-Nov-99
		GB_HTG2:HS500L14	164856	AL023583	Homo sapiens chromosome 6 clone RP3-500L14 map p23-24.3, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	34,632	30-Nov-99
		GB_HTG2:HS500L14	164856	AL023583	Homo sapiens chromosome 6 clone RP3-500L14 map p23-24.3, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Homo sapiens	34,632	30-Nov-99
rx01040	1026	GB_EST24:AI193549	479	AI193549	qe70e06.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:1744354 3', mRNA sequence.	Homo sapiens	40,126	29-OCT-1998
		GB_PR2:HSU38545	3609	U38545	Human ARF-activated phosphatidylincholine-specific phospholipase D1a (hPLD1) mRNA, complete cds.	Homo sapiens	38,652	10-MAR-1997
		GB_PR2:AC002481	28244	AC002481	Human cosmid clone LUCA12 from 3p21.3, complete sequence.	Homo sapiens	39,643	21-Aug-97

TABLE 4: ALIGNMENT RESULTS

rx01041	276	GB_HTG6:AC007957 GB_PR2:AP000552 GB_PR3:HS57A13	212658 157086 169693	AC007957 AP000552 Z83848	Homo sapiens, *** SEQUENCING IN PROGRESS *** Homo sapiens genomic DNA, chromosome 22q11.2, BCL2L2 region, clone:KB1183D5. Human DNA sequence from PAC 57A13 between markers DXS6791 and DXS6038 on chromosome X contains glutamate receptor subunit GluRc, ESTs, STS and polymorphic CA repeat.	Homo sapiens Homo sapiens Homo sapiens	40,809 40,809 37,647	26-Nov-99 01-OCT-1999 23-Nov-99
rx01042	1401	GB_BA1:MTCY10G2	38970	Z92539	Mycobacterium tuberculosis H37Rv complete genome; segment 47/162.	Mycobacterium tuberculosis	36,023	17-Jun-98
		GB_BA1:MTCY10G2	38970	Z92539	Mycobacterium tuberculosis H37Rv complete genome; segment 47/162.	Mycobacterium tuberculosis	37,010	17-Jun-98
		GB_EST28:AS51042	538	AI551042	vx33d11.x1 Stragene mouse lung 937302 Mus musculus cDNA clone IMAGE:1277013 3', mRNA sequence.	Mus musculus	38,806	23-MAR-1999
rx01043	696	GB_BA1:AF006658 GB_BA1:MLB1790G	2500 37617	AF006658 Z14314	Bacteroides fragilis beta-glucosidase gene, complete cds. M.leprae genes rplL, rpoB, rpoC, end, rpsL, rpsG, efg, tuf, rpsJ, rplC for ribosomal protein L7, RNA polymerase beta subunit, RNA polymerase beta' subunit, endonuclease, ribosomal protein S7, ribosomal protein S12, elongation factor G, elongation factor Tu, ribosomal protein S10, ribosomal protein L3 and mkl gene.	Bacteroides fragilis Mycobacterium leprae	39,156 39,970	12-Jul-97 11-Feb-93
rx01044	1380	GB_BA1:AF006658 GB_HTG6:AC010998 GB_HTG6:AC010998 GB_BA2:AE000939	2500 144338 144338 10599	AF006658 AC010998 AC010998 AE000939	Bacteroides fragilis beta-glucosidase gene, complete cds. Homo sapiens clone RP11-9516, *** SEQUENCING IN PROGRESS *** Homo sapiens clone RP11-9516, *** SEQUENCING IN PROGRESS *** Methanobacterium thermoautotrophicum from bases 1698671 to 1709269 (section 145 of 148) of the complete genome.	Bacteroides fragilis Homo sapiens Homo sapiens Methanobacterium thermoautotrophicum	36,472 37,630 37,864 34,480	12-Jul-97 08-DEC-1999 08-DEC-1999 15-Nov-97
rx01045	1947	GB_VI:FCVCF6A GB_OM:CATFLVPOI	8440 3639	M18247 L06140	Feline leukemia virus, subgroup A (FeLV-FAIDS), complete nucleotide sequence. Felis catus endogenous FeLV proviral polyprotein (protease (PRO), reverse transcriptase (RT), integrase/endonuclease (INT)) and pol pseudogene, 3' end.	Feline leukemia virus Felis catus	37,617 41,966	29-MAR-1996 21-Aug-95
		GB_VI:CEAVCG	9189	M33677	Caprine arthritis encephalitis virus, complete proviral genome.	Caprine arthritis-encephalitis virus	36,297	04-MAR-1996
rx01046	1902	GB_HTG3:AC008423	177734	AC008423	Homo sapiens chromosome 5 clone CIT-HSPC_298N6, *** SEQUENCING IN PROGRESS *** unordered pieces.	Homo sapiens	38,720	3-Aug-99
		GB_HTG3:AC008423	177734	AC008423	Homo sapiens chromosome 5 clone CIT-HSPC_298N6, *** SEQUENCING IN PROGRESS *** unordered pieces.	Homo sapiens	38,720	3-Aug-99
		GB_HTG3:AC008423	177734	AC008423	Homo sapiens chromosome 5 clone CIT-HSPC_298N6, *** SEQUENCING IN PROGRESS *** unordered pieces.	Homo sapiens	35,882	3-Aug-99
rx01047	597	GB_EST20:AA842685	510	AA842685	MBAFcZ9C11T3 Brugia malayi adult female cDNA (SAW96MLW-BmAF) Brugia malayi cDNA clone AFCZ9C11 5', mRNA sequence.	Brugia malayi	37,965	02-MAR-1998
		GB_EST20:AA842685	510	AA842685	MBAFcZ9C11T3 Brugia malayi adult female cDNA (SAW96MLW-BmAF) Brugia malayi cDNA clone AFCZ9C11 5', mRNA sequence.	Brugia malayi	41,697	02-MAR-1998
rx01058	444	GB_GSS9:AQ160800	745	AQ160800	nbxb0006C07r CUGI Rice BAC Library Oryza sativa genomic clone nbxb0006C07r, genomic survey sequence.	Oryza sativa	38,242	12-Sep-98
		GB_GSS3:B10162	1102	B10162	F11B10-Sp6 IGF Arabidopsis thaliana genomic clone F11B10, genomic survey sequence.	Arabidopsis thaliana	42,263	14-MAY-1997
		GB_BA1:AB032799	9077	AB032799	Chromobacterium violaceum violacein biosynthetic gene cluster (vioA, vio B, vio C, vio D), complete cds.	Chromobacterium violaceum	34,475	02-OCT-1999
rx01063	453	GB_GSS4:AQ707752	510	AQ707752	HS_5560_A2_G07_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1136 Col=14 Row=M, genomic survey sequence.	Homo sapiens	36,932	7-Jul-99
		GB_GSS4:AQ707752	510	AQ707752	HS_5560_A2_G07_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1136 Col=14 Row=M, genomic survey sequence.	Homo sapiens	35,885	7-Jul-99
rx01066	849	GB_BA2:U32709	10010	U32709	Haemophilus influenzae Rd section 24 of 163 of the complete genome.	Haemophilus influenzae Rd	36,158	29-MAY-1998
		GB_RO:AB009615	1515	AB009615	Mus musculus mRNA for type II phosphatidylinositolphosphate kinase-alpha, complete cds.	Mus musculus	37,861	13-Feb-99

TABLE 4: ALIGNMENT RESULTS

GB_RO:AB032899	1914	AB032899	Rattus norvegicus PIPK2 alpha mRNA for phosphatidylinositol 5-phosphate 4-kinase alpha. complete cds.	Rattus norvegicus	38,480	07-OCT-1999
GB_HTG4:AC006091	176878	AC006091	Drosophila melanogaster chromosome 3 clone BACR48G05 (D475) RPCI-98 48 G.5 map 91F1-91F13 strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 4 unordered pieces.	Drosophila melanogaster	35,539	27-OCT-1999
GB_HTG4:AC006091	176878	AC006091	Drosophila melanogaster chromosome 3 clone BACR48G05 (D475) RPCI-98 48 G.5 map 91F1-91F13 strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 4 unordered pieces.	Drosophila melanogaster	35,539	27-OCT-1999
GB_HTG2:AC008141	100729	AC008141	Drosophila melanogaster chromosome 3 clone BACR17F04 (D988) RPCI-98 17 F.4 map 91F-91F strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 69 unordered pieces.	Drosophila melanogaster	34,415	2-Aug-99
GB_EST15:AA531901	524	AA531901	TgESTz32g09.r1 TgME49 invivo Bradyzoite cDNA size selected Toxoplasma gondii cDNA clone tgz32g09.r1 5' mRNA sequence.	Toxoplasma gondii	43,005	22-Jul-97
GB_EST15:AA520183	527	AA520183	TgESTz39d01.s1 TgME49 invivo Bradyzoite cDNA size selected Toxoplasma gondii cDNA clone tgz339d01.s1 3' mRNA sequence.	Toxoplasma gondii	40,664	16-Jul-97
GB_HTG6:AC010846	187611	AC010846	Drosophila melanogaster chromosome X clone BACR13G13 (D894) RPCI-98 13 G.13 map 14B-14C strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 97 unordered pieces.	Drosophila melanogaster	36,679	03-DEC-1999
GB_EST20:AA880319	450	AA880319	vx39h01.r1 Stratiene mouse lung 937302 Mus musculus cDNA clone IMAGE:1277617 5' mRNA sequence.	Mus musculus	40,724	26-MAR-1998
GB_GSS14:AQ558382	435	AQ558382	HS_2068_B1_F06_TTC C1T Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2068 Col=11 Row=L genomic survey sequence.	Homo sapiens	36,882	29-MAY-1999
GB_GSS15:AQ600385	483	AQ600385	HS_5357_B2_C05_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=933 Col=10 Row=F genomic survey sequence.	Homo sapiens	40,476	10-Jun-99
GB_BA1:PSEHEDDH	3060	M74256	Pseudomonas aeruginosa 6-phosphogluconate dehydratase (edd) gene, and glyceraldehyde-3-phosphate dehydrogenase (gap) gene, complete cds.	Pseudomonas aeruginosa	39,657	30-Nov-93
GB_BA1:CGL007732	4460	AJ007732	Corynebacterium glutamicum 3' ppc gene, secG gene, amt gene, ocd gene and 5' soxA gene.	Corynebacterium glutamicum	39,168	7-Jan-99
GB_EST9:AA066016	406	AA066016	mi52f12.r1 Stratiene mouse testis (#937308) Mus musculus cDNA clone IMAGE:515663 5' mRNA sequence.	Mus musculus	43,382	3-Feb-97
GB_EST21:AA986543	445	AA986543	ue14f08.x1 Sugano mouse embryo mewa Mus musculus cDNA clone IMAGE:1480359 3' mRNA sequence.	Mus musculus	31,236	28-MAY-1998
GB_EST22:A035794	509	A035794	ue17d01.y1 Sugano mouse embryo mewa Mus musculus cDNA clone IMAGE:1480609 5' mRNA sequence.	Mus musculus	42,264	26-Jun-98
GB_EST22:A006506	384	A006506	ue14f08.y1 Sugano mouse embryo mewa Mus musculus cDNA clone IMAGE:1480359 5' mRNA sequence.	Mus musculus	46,637	12-Jun-98
GB_HTG2:AC007741	162450	AC007741	Homo sapiens clone NH0340F16. *** SEQUENCING IN PROGRESS ***. 3 unordered pieces.	Homo sapiens	38,209	5-Jun-99
GB_HTG2:AC007741	162450	AC007741	Homo sapiens clone NH0340F16. *** SEQUENCING IN PROGRESS ***. 3 unordered pieces.	Homo sapiens	38,209	5-Jun-99
GB_EST33:AV072325	317	AV072325	AV072325 Mus musculus stomach C57BL/6J adult Mus musculus cDNA clone 2200003E03, mRNA sequence.	Mus musculus	48,485	24-Jun-99
GB_BA2:RCPHSYNG	45959	Z11165	R.capsulatus complete photosynthesis gene cluster.	Rhodobacter capsulatus	36,603	2-Sep-99
GB_BA2:RCPHSYNG	45959	Z11165	R.capsulatus complete photosynthesis gene cluster.	Rhodobacter capsulatus	37,989	2-Sep-99
GB_PR4:AF073931	7898	AF073931	Homo sapiens low-voltage activated calcium channel alpha 1H mRNA, complete cds.	Homo sapiens	37,953	04-MAR-1999
GB_BA2:AF112535	4363	AF112535	Corynebacterium glutamicum putative glutaredoxin NrdH (nrdH), and ribonucleotide reductase alpha-chain (nrdE) genes, complete cds.	Corynebacterium glutamicum	99,499	5-Aug-99
GB_PR3:HS3D2	1789	AF053138	Homo sapiens histone deacetylase 3 gene, exons 11, 12, 13 and partial cds.	Homo sapiens	33,512	28-MAR-1998
GB_PR4:AF059650	16015	AF059650	Homo sapiens histone deacetylase 3 (HDAC3) gene, complete cds.	Homo sapiens	38,814	03-MAR-1999
GB_EST4:H55032	951	H55032	HHU58a Sorghum bicolor cv. TX430 Sorghum bicolor cDNA clone HHU58 5' similar to transketolase, chloroplast (TKL1), mRNA sequence.	Sorghum bicolor	41,111	27-Sep-99
GB_HTG2:HSBA27F12	123489	AL109914	Homo sapiens chromosome 6 clone RP11-27F12 map p22 3-24. *** SEQUENCING IN PROGRESS ***. in unordered pieces.	Homo sapiens	35,156	30-Nov-99
GB_HTG2:HSBA27F12	123489	AL109914	Homo sapiens chromosome 6 clone RP11-27F12 map p22 3-24. *** SEQUENCING IN PROGRESS ***. in unordered pieces.	Homo sapiens	35,156	30-Nov-99
GB_HTG5:AC010202	170004	AC010202	Homo sapiens chromosome 12q seeders clone RP11-210L7. *** SEQUENCING IN PROGRESS ***. 40 unordered pieces.	Homo sapiens	37,313	6-Nov-99



TABLE 4: ALIGNMENT RESULTS

GB_HTG5:AC010202	170004	AC010202	Homo sapiens chromosome 12q seeders clone RP11-210L7, *** SEQUENCING IN PROGRESS ***, 40 unordered pieces.	Homo sapiens	37,422	6-Nov-99
GB_PR1:HSIGFACI	7260	X57025	Homo IGF-I mRNA for insulin-like growth factor I.	Homo sapiens	38,043	17-Feb-92
GB_BA1:ECORELA	4034	J04039	E. coli relA gene encoding ATP-GTP 3'-pyrophosphotransferase, complete cds.	Escherichia coli	54,711	16-Nov-93
GB_BA2:ECU29580	13234	U29580	Escherichia coli K-12 genome; approximately 62 minute region.	Escherichia coli	37,327	5-Apr-99
GB_BA2:AE000362	12595	AE000362	Escherichia coli K-12 MG1655 section 252 of 400 of the complete genome.	Escherichia coli	37,327	12-Nov-98
GB_PL1:PCX24CRY	357	Z34459	P. cryptogaea X24 gene for cryptogaein.	Phytophthora cryptogaea	43,310	19-Sep-96
GB_EST2:AI244520	414	AI244520	qk14c08.x1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1868942 3', mRNA sequence.	Homo sapiens	33,528	28-Jan-99
GB_RO:MM26SPROT	1479	Y13071	Mus musculus mRNA for 26S proteasome non-ATPase subunit.	Mus musculus	37,941	10-Sep-98
GB_HTG3:AC009219	127519	AC009219	Drosophila melanogaster chromosome 3 clone BACR32N16 (D973) RPCL98 32.N.16 map 86C-86C strain Y; cn bw sp. *** SEQUENCING IN PROGRESS***, 74 unordered pieces.	Drosophila melanogaster	35,080	20-Aug-99
GB_HTG3:AC009219	127519	AC009219	Drosophila melanogaster chromosome 3 clone BACR32N16 (D973) RPCL98 32.N.16 map 86C-86C strain Y; cn bw sp. *** SEQUENCING IN PROGRESS***, 74 unordered pieces.	Drosophila melanogaster	35,080	20-Aug-99
GB_PR4:AC006065	191134	AC006065	Homo sapiens 12q24.2 BAC RPC111-407A16 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	Homo sapiens	37,453	27-Feb-99
GB_EST13:AA340958	338	AA340958	EST46332 Fetal kidney II Homo sapiens cDNA 5' end, mRNA sequence.	Homo sapiens	36,596	21-Apr-97
GB_RO:MUSBCL22	5806	L31532	Mus musculus bc1-2 alpha gene, exon 2.	Mus musculus	33,913	5-Apr-94
GB_BA2:AE001165	13021	AE001165	Borrelia burgdorferi (section 51 of 70) of the complete genome.	Borrelia burgdorferi	31,412	15-DEC-1997
GB_HTG1:HS1030M6	173804	AL035089	Homo sapiens chromosome 20 clone RP5-1030M6, *** SEQUENCING IN PROGRESS***, in unordered pieces.	Homo sapiens	34,935	23-Nov-99
GB_HTG1:HS1030M6	173804	AL035089	Homo sapiens chromosome 20 clone RP5-1030M6, *** SEQUENCING IN PROGRESS***, in unordered pieces.	Homo sapiens	34,935	23-Nov-99
GB_HTG2:AC005057	99370	AC005057	Homo sapiens clone RG052HB06, *** SEQUENCING IN PROGRESS***, 11 unordered pieces.	Homo sapiens	35,897	12-Jun-98
GB_PR2:HS193N13	122961	AL078600	Human DNA sequence from clone RP1-193N13 on chromosome 6q21-22.31, complete sequence.	Homo sapiens	37,115	22-Nov-99
GB_EST34:AV139054	287	AV139054	AV139054 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA clone 2810048D09, mRNA sequence.	Mus musculus	35,540	1-Jul-99
GB_PL1:GTU21246	926	U21246	Gracilariopsis tenuifrons internal transcribed spacer region of the ribosomal repeat, ITS1, 5.8S rRNA Gracilariopsis tenuifrons gene and ITS2, complete sequence.	Gracilariopsis tenuifrons	37,200	12-MAR-1995
GB_BA1:AB003332	1424	AB003332	Anabaena variabilis rbpF gene for RNA binding protein, complete cds.	Anabaena variabilis	35,958	21-MAY-1999
GB_BA1:BPETOXOP	9342	L10720	Bordetella pertussis toxin liberation operon.	Bordetella pertussis	40,107	9-Jul-93
GB_PAT:IS0844	951	I50844	Sequence 12 from patent US 5643747.	Unknown.	39,973	07-OCT-1997
GB_GSS3:B36708	438	B36708	HS-1041-B1-C05-MF.abi CIT Human Genomic Sperm Library C Homo sapiens genomic clone Plate=CT 823 Col=9 Row=F, genomic survey sequence.	Homo sapiens	37,300	17-OCT-1997
GB_PL2:F11A17	102077	AC007932	Arabidopsis thaliana chromosome 1 BAC F11A17 sequence, complete sequence.	Arabidopsis thaliana	37,488	16-Aug-99
GB_EST22:AI043264	283	AI043264	TENU0904 T. cruzi epimastigote normalized cDNA Library Trypanosoma cruzi cDNA clone 2114 3', mRNA sequence.	Trypanosoma cruzi	40,989	1-Jul-98
GB_PAT:I78756	737	I78756	Sequence 12 from patent US 5693781.	Unknown.	40,975	3-Apr-98
GB_PAT:I92045	737	I92045	Sequence 12 from patent US 5726299.	Unknown.	40,975	01-DEC-1998
GB_PL1:MTAPCG	100314	X55026	P. anserina complete mitochondrial genome.	Mitochondrion Podospora anserina	34,477	08-DEC-1997
GB_HTG4:AC010148	228794	AC010148	Homo sapiens chromosome unknown clone NH0367B19, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	42,130	29-OCT-1999
GB_HTG4:AC010148	228794	AC010148	Homo sapiens chromosome unknown clone NH0367B19, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	42,130	29-OCT-1999
GB_HTG4:AC010148	228794	AC010148	Homo sapiens chromosome unknown clone NH0367B19, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	37,559	29-OCT-1999
GB_EST33:AV090612	274	AV090612	AV090612 Mus musculus tongue C57BL/6J adult Mus musculus cDNA clone 2310051C21, mRNA sequence.	Mus musculus	33,212	28-Jun-99
GB_EST33:AV090612	274	AV090612	AV090612 Mus musculus tongue C57BL/6J adult Mus musculus cDNA clone 2310051C21, mRNA sequence.	Mus musculus	34,529	28-Jun-99

TABLE 4: ALIGNMENT RESULTS

rx01127	1143	GB_BA1:CGLEUB	2042	Y09578	C. glutamicum leuB gene.	Corynebacterium glutamicum	99,913	02-MAR-1999
		GB_BA1:MTV012	70287	AL021287	Mycobacterium tuberculosis H37Rv complete genome; segment 132/162.	Mycobacterium tuberculosis	39,295	23-Jun-99
rx01128	1137	GB_BA1:SC1C2	42210	AL031124	Streptomyces coelicolor cosmid 1C2.	Streptomyces coelicolor	62,456	15-Jan-99
		GB_BA1:CGLEUB	2042	Y09578	C. glutamicum leuB gene.	Corynebacterium glutamicum	38,515	02-MAR-1999
		GB_GSS8:AO066341	241	AQ066341	HS_2243_B1_A02_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2243 Col=3 Row=B, genomic survey sequence.	Homo sapiens	39,004	4-Aug-98
		GB_EST24:A168493	459	A168493	ou64g08.s1 NCI CGAP_Br2 Homo sapiens cDNA clone IMAGE:1632638 3' similar to gb:D90209 DNA-BINDING PROTEIN TAXREB67 (HUMAN); mRNA sequence.	Homo sapiens	42,117	23-OCT-1998
rx01129	1989	GB_BA1:MTCI28	36300	Z97050	Mycobacterium tuberculosis H37Rv complete genome; segment 10/162.	Mycobacterium tuberculosis	38,159	23-Jun-98
		GB_HTG1:LMFL6852	37286	AL034359	Leishmania major chromosome 4 clone L6852 strain Freidlin, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Leishmania major	38,485	29-Apr-99
		GB_HTG1:LMFL6852	37286	AL034359	Leishmania major chromosome 4 clone L6852 strain Freidlin, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Leishmania major	38,485	29-Apr-99
rx01131	990	GB_BA1:CGLEUB	2042	Y09578	C. glutamicum leuB gene.	Corynebacterium glutamicum	99,313	02-MAR-1999
		GB_PR4:AC007253	225699	AC007253	Homo sapiens BAC clone NH0454P05 from 2, complete sequence.	Homo sapiens	35,421	22-OCT-1999
		GB_HTG3:AC011305	171067	AC011305	Homo sapiens clone NH0390E09, *** SEQUENCING IN PROGRESS ***; 1 unordered pieces.	Homo sapiens	35,955	05-OCT-1999
rx01134	871	GB_HTG4:AC009375	137069	AC009375	Drosophila melanogaster chromosome 3L75A1 clone RPC198-44L18, *** SEQUENCING IN PROGRESS ***; 59 unordered pieces.	Drosophila melanogaster	37,380	16-OCT-1999
		GB_HTG4:AC009375	137069	AC009375	Drosophila melanogaster chromosome 3L75A1 clone RPC198-44L18, *** SEQUENCING IN PROGRESS ***; 59 unordered pieces.	Drosophila melanogaster	37,380	16-OCT-1999
		GB_HTG4:AC009375	137069	AC009375	Drosophila melanogaster chromosome 3L75A1 clone RPC198-44L18, *** SEQUENCING IN PROGRESS ***; 59 unordered pieces.	Drosophila melanogaster	37,380	16-OCT-1999
rx01137	483	GB_HTG4:AC009375	137069	AC009375	Drosophila melanogaster chromosome 3L75A1 clone RPC198-44L18, *** SEQUENCING IN PROGRESS ***; 59 unordered pieces.	Drosophila melanogaster	37,380	16-OCT-1999
		GB_HTG4:AC009375	137069	AC009375	Drosophila melanogaster chromosome 3L75A1 clone RPC198-44L18, *** SEQUENCING IN PROGRESS ***; 59 unordered pieces.	Drosophila melanogaster	37,380	16-OCT-1999
rx01140	1056	GB_IN2:AC005421	69992	AC005421	Drosophila melanogaster, chromosome 2L, region 22A1-22A1, P1 clone DS03601, complete sequence.	Drosophila melanogaster	33,333	31-OCT-1998
		GB_GSS1:CNS00KX9	1101	AL078350	Drosophila melanogaster genome survey sequence TET3 end of BAC:BACR23A23 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	Drosophila melanogaster	30,670	3-Jun-99
		GB_IN2:AC005421	69992	AC005421	Drosophila melanogaster, chromosome 2L, region 22A1-22A1, P1 clone DS03601, complete sequence.	Drosophila melanogaster	33,333	31-OCT-1998
rx01140	1056	GB_PR4:AC007948	99904	AC007948	Genomic sequence of Homo sapiens clone R417F14A from chromosome 18, complete sequence.	Homo sapiens	36,249	6-Nov-99
		GB_HTG3:AC006278	140290	AC006278	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***; 18 unordered pieces.	Plasmodium falciparum	36,592	23-Sep-99
		GB_HTG3:AC006278	140290	AC006278	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***; 18 unordered pieces.	Plasmodium falciparum	36,592	23-Sep-99
rx01148	723	GB_BA1:MTCY261	27322	Z97559	Mycobacterium tuberculosis H37Rv complete genome; segment 95/162.	Mycobacterium tuberculosis	50,139	17-Jun-98
		GB_EST15:AA501229	548	AA501229	vh62g12.r1 Knowles Solter mouse 11 5day limb bud Mus musculus cDNA clone IMAGE:891622 5' similar to TR:G762951 G762951 PRPL-2 PROTEIN; mRNA sequence.	Mus musculus	36,250	1-Jul-97
		GB_PR4:AF106062	1306	AF106062	Homo sapiens Wiskott-Aldrich syndrome protein interacting protein (WASPIP) mRNA, partial cds.	Homo sapiens	43,205	31-Jul-99
rx01153	543	GB_BA1:MTCY261	27322	Z97559	Mycobacterium tuberculosis H37Rv complete genome; segment 95/162.	Mycobacterium tuberculosis	40,000	17-Jun-98
		GB_BA1:MLCB2533	40245	AL035310	Mycobacterium leprae cosmid B2533.	Mycobacterium leprae	61,765	27-Aug-99
		GB_BA1:U00017	42157	U00017	Mycobacterium leprae cosmid B2126.	Mycobacterium leprae	39,615	01-MAR-1994

TABLE 4: ALIGNMENT RESULTS

rx01154	677	GB_BA1:MTCY49	39430	Z73966	Mycobacterium tuberculosis H37Rv complete genome; segment 93/162.	Mycobacterium tuberculosis	37,615	24-Jun-99
		GB_BA1:U00017	42157	U00017	Mycobacterium leprae cosmid B2126.	Mycobacterium leprae	36,957	01-MAR-1994
		GB_BA2:AF086832	10612	AF086832	Streptomyces coelicolor putative ferredoxin, ARC (arc), 20S proteasome beta-subunit precursor (prcB), 20S proteasome alpha-subunit (prcA), putative LacI family repressor, and putative transporter genes, complete cds; and unknown genes.	Streptomyces coelicolor	53,354	30-Jun-99
rx01155	1570	GB_BA1:MTCY49	39430	Z73966	Mycobacterium tuberculosis H37Rv complete genome; segment 93/162.	Mycobacterium tuberculosis	37,025	24-Jun-99
		GB_BA1:REZ82005	3301	Z82005	R. erythropolis DNA, 20S proteasome structural genes region (3301 bp).	Rhodococcus erythropolis	54,128	8-Nov-97
		GB_BA2:RSU26422	3554	U26422	Rhodococcus erythropolis ORF6(2), ORF7(2), proteasome beta-type subunit 2 (prcB(2)), and proteasome alpha-type subunit 2 (prcA(2)) genes, complete cds.	Rhodococcus erythropolis	45,951	08-OCT-1997
rx01156	257	GB_BA2:RSU26422	3554	U26422	Rhodococcus erythropolis ORF6(2), ORF7(2), proteasome beta-type subunit 2 (prcB(2)), and proteasome alpha-type subunit 2 (prcA(2)) genes, complete cds.	Rhodococcus erythropolis	54,724	08-OCT-1997
		GB_HTG6:AC009218	126649	AC009218	Drosophila melanogaster chromosome 2 clone BACR33D17 (D945) RPCI-98 33.D.17 map 57B-57B strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 76 unordered pieces.	Drosophila melanogaster	42,570	02-DEC-1999
rx01158	1065	GB_OV:QULNFW	7933	D13223	Japanese quail genomic DNA for neurofilament-L (NF-L).	Coturnix coturnix	44,758	3-Feb-99
		GB_HTG3:AC011401	321277	AC011401	Homo sapiens chromosome 5 clone CIT978SKB_35K5, *** SEQUENCING IN PROGRESS ***. 65 unordered pieces.	Homo sapiens	35,048	06-OCT-1999
		GB_HTG3:AC011401	321277	AC011401	Homo sapiens chromosome 5 clone CIT978SKB_35K5, *** SEQUENCING IN PROGRESS ***. 65 unordered pieces.	Homo sapiens	35,048	06-OCT-1999
rx01159	438	GB_PR3:AC004386	172657	AC004386	Homo Sapiens Chromosome X clone BWXD691, complete sequence.	Homo sapiens	37,223	10-Apr-98
		GB_PAT:192035	413	I92035	Sequence 2 from patent US 5726299.	Unknown.	78,169	01-DEC-1998
		GB_PAT:178746	413	I78746	Sequence 2 from patent US 5693781.	Unknown.	78,169	3-Apr-98
		GB_HTG3:AC009209	108370	AC009209	Drosophila melanogaster chromosome 2 clone BACR24G16 (D1051) RPCI-98 24.G.16 map 47D-47D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 93 unordered pieces.	Drosophila melanogaster	43,590	20-Aug-99
rx01160	998	GB_PAT:178746	413	I78746	Sequence 2 from patent US 5693781.	Unknown.	96,032	3-Apr-98
		GB_PAT:192035	413	I92035	Sequence 2 from patent US 5726299.	Unknown.	96,032	01-DEC-1998
		GB_BA2:AF014804	6449	AF014804	Neisseria meningitidis PglB (pglB), PglC (pglC), and AvtA (avtA) genes, complete cds.	Neisseria meningitidis	37,977	3-Sep-99
rx01165	696	GB_HTG2:AC007851	128979	AC007851	Drosophila melanogaster chromosome 2 clone BACR08M19 (D615) RPCI-98 06.M.19 map 50C-50D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 86 unordered pieces.	Drosophila melanogaster	37,900	2-Aug-99
		GB_HTG2:AC007851	128979	AC007851	Drosophila melanogaster chromosome 2 clone BACR08M19 (D615) RPCI-98 06.M.19 map 50C-50D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 86 unordered pieces.	Drosophila melanogaster	37,900	2-Aug-99
		GB_HTG2:AC007851	128979	AC007851	Drosophila melanogaster chromosome 2 clone BACR08M19 (D615) RPCI-98 06.M.19 map 50C-50D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 86 unordered pieces.	Drosophila melanogaster	34,114	2-Aug-99
rx01166	1428	GB_BA1:LEUG6PD	1957	M64446	L. mesenteroides glucose-6-phosphate dehydrogenase gene, complete cds.	Leuconostoc mesenteroides	41,259	26-Apr-93
		GB_HTG3:AF188026	101456	AF188026	Homo sapiens chromosome 8 clone BAC 2379L20 map 8q24, *** SEQUENCING IN PROGRESS ***. In ordered pieces.	Homo sapiens	35,535	08-OCT-1999
		GB_HTG3:AF188026	101456	AF188026	Homo sapiens chromosome 8 clone BAC 2379L20 map 8q24, *** SEQUENCING IN PROGRESS ***. In ordered pieces.	Homo sapiens	35,535	08-OCT-1999
rx01167	519	GB_PR2:HS1026E2	100418	AL022143	Human DNA sequence from clone 1026E2 on chromosome 1q24.1-25.3 EST, CA repeat. STS, GSS, complete sequence.	Homo sapiens	38,281	23-Nov-99
		GB_EST36:A1900015	496	A1900015	sb9705.y1 Gm-c1012 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1012-634 5' similar to WP.T11G6.8 CE08432 RNA RECOGNITION MOTIF. mRNA sequence.	Glycine max	42,045	06-DEC-1999
		GB_PR2:HS1026E2	100418	AL022143	Human DNA sequence from clone 1026E2 on chromosome 1q24.1-25.3 EST, CA repeat. STS, GSS, complete sequence.	Homo sapiens	40,990	23-Nov-99
rx01169	1119	GB_HTG2:AC005978	92586	AC005978	Drosophila melanogaster chromosome 2 clone DS00678 (D449) map 59D3-59D4 strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 6 unordered pieces.	Drosophila melanogaster	34,266	30-Jul-99
		GB_HTG2:AC005978	92586	AC005978	Drosophila melanogaster chromosome 2 clone DS00678 (D449) map 59D3-59D4 strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 6 unordered pieces.	Drosophila melanogaster	34,266	30-Jul-99

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

GB_HTG3:AC008304	91552	AC008304	Drosophila melanogaster chromosome 2 clone BACR04G19 (D646) RPCI-98 04.G.19 map 59D2-59D3 strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 90 unordered pieces.	31,278	Drosophila melanogaster	20-Sep-99
GB_PR1:AB014524	6542	AB014524	Homo sapiens mRNA for KIAA0624 protein, partial cds.	43,478	Homo sapiens	6-Feb-99
GB_EST38:AW016078	496	AW016078	UI-H-BIOP-abf-h-01-U-I.s1 NCI_CGAP_Sub2 Homo sapiens cDNA clone IMAGE:2711665 3', mRNA sequence.	36,667	Homo sapiens	10-Sep-99
GB_EST24:AI193238	323	AI193238	qe56c06.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:1742986 3', mRNA sequence.	37,427	Homo sapiens	29-OCT-1998
GB_HTG2:AC007548	110249	AC007548	Drosophila melanogaster chromosome 2 clone BACR48M17 (D614) RPCI-98 48.M.17 map 41C-41D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 66 unordered pieces.	35,634	Drosophila melanogaster	2-Aug-99
GB_PR3:HS273F20	111253	AL034371	Human DNA sequence from clone 273F20 on chromosome 6q16.1-16.3 Contains ESTs, STSs and GSSs, complete sequence.	34,582	Homo sapiens	23-Nov-99
GB_HTG2:AC007548	110249	AC007548	Drosophila melanogaster chromosome 2 clone BACR48M17 (D614) RPCI-98 48.M.17 map 41C-41D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 66 unordered pieces.	35,634	Drosophila melanogaster	2-Aug-99
GB_GSS8:B93272	338	B93272	CIT-HSP-2171E14.TF CIT-HSP Homo sapiens genomic clone 2171E14, genomic survey sequence.	42,012	Homo sapiens	25-Jun-98
GB_EST35:AV156265	282	AV156265	AV156265 Mus musculus head C57BL/6J 12-day embryo Mus musculus cDNA clone 3000001103, mRNA sequence.	41,350	Mus musculus	7-Jul-99
GB_ROAF035777	2154	AF035777	Mus musculus somatostatin receptor subtype 5 (sst5) gene, complete cds.	38,928	Mus musculus	05-DEC-1997
GB_HTG1:CEY51A2_3	110000	Z99275	Caenorhabditis elegans chromosome V clone Y51A2, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	38,045	Caenorhabditis elegans	3-Dec-98
GB_IN1:CEY51A2D	139259	AL021497	Caenorhabditis elegans cosmid Y51A2D, complete sequence.	38,045	Caenorhabditis elegans	2-Sep-99
GB_HTG1:CEY51A2_3	110000	Z99275	Caenorhabditis elegans chromosome V clone Y51A2, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	38,045	Caenorhabditis elegans	3-Dec-98
GB_BA1:MTV021	23400	AL021957	Mycobacterium tuberculosis H37Rv complete genome; segment 97/162.	39,159	Mycobacterium tuberculosis	18-Jun-98
GB_BA1:MSGB1551CS	36548	L78813	Mycobacterium leprae cosmid B1551 DNA sequence.	53,215	Mycobacterium leprae	15-Jun-96
GB_BA1:MSGB1554CS	36548	L78814	Mycobacterium leprae cosmid B1554 DNA sequence.	53,215	Mycobacterium leprae	15-Jun-96
GB_PR4:AC006126	46100	AC006126	Homo sapiens chromosome 19, cosmid F18718, complete sequence.	36,832	Homo sapiens	17-DEC-1998
GB_PR4:AC006126	46100	AC006126	Homo sapiens chromosome 19, cosmid F18718, complete sequence.	40,463	Homo sapiens	17-DEC-1998
GB_EST23:AI128078	443	AI128078	qc47d10.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone IMAGE:1712755 3' similar to SW/IF16_HUMAN Q16666 GAMMA-INTERFERON-INDUCIBLE PROTEIN IFI-16, mRNA sequence.	38,318	Homo sapiens	27-OCT-1998
GB_HTG2:HSG248A21	96783	AL118512	Homo sapiens chromosome 1 clone GS1-248A21, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	36,994	Homo sapiens	30-Nov-99
GB_HTG2:HSG248A21	96783	AL118512	Homo sapiens chromosome 1 clone GS1-248A21, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	36,994	Homo sapiens	30-Nov-99
GB_PR3:AC004237	38715	AC004237	Homo sapiens chromosome 5, P1 clone 565a12 (LBNL H23), complete sequence.	36,514	Homo sapiens	27-Feb-98
GB_BA1:MTCY274	39991	Z74024	Mycobacterium tuberculosis H37Rv complete genome; segment 126/162.	34,758	Mycobacterium tuberculosis	19-Jun-98
GB_BA1:SC6A9	39461	AL031035	Streptomyces coelicolor cosmid 6A9	38,500	Streptomyces coelicolor	24-Jul-98
GB_BA2:AE000689	14698	AE000689	Aquifex aeolicus section 21 of 109 of the complete genome.	45,379	Aquifex aeolicus	25-MAR-1998
GB_EST5:H60943	361	H60943	Yr14g08.r1 Soares_fetal_liver_spleen_1NFLS Homo sapiens cDNA clone IMAGE:205310 5', mRNA sequence.	39,205	Homo sapiens	06-OCT-1995
GB_GSS1:AG018858	570	AG018858	Homo sapiens genomic DNA, 21q region, clone: B125C11 SpN045(-21), genomic survey sequence.	41,423	Homo sapiens	10-OCT-1999
GB_GSS13:AQ454839	505	AQ454839	HS_5218_A1_E03_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=794 Col=5 Row=I, genomic survey sequence.	34,599	Homo sapiens	21-Apr-99
GB_PR3:AF022813	1358	AF022813	Homo sapiens tetraspan (NAG-2) mRNA, complete cds.	40,614	Homo sapiens	18-Nov-97

**TABLE 4: ALIGNMENT RESULTS**

GB_EST34:AI798721	765	AI798721	we91g10.x1 Soares_NFL_I_GBC_S1 Homo sapiens cDNA clone IMAGE:2348514 3' similar to SW:NAG2_HUMAN O14817 NOVEL ANTIGEN 2, mRNA sequence.	Homo sapiens	35,256	6-Jul-99
GB_EST34:AI808898	696	AI808898	wf66d02.x1 Soares_NFL_I_GBC_S1 Homo sapiens cDNA clone IMAGE:2360547 3' similar to SW:NAG2_HUMAN O14817 NOVEL ANTIGEN 2, mRNA sequence.	Homo sapiens	35,354	7-Jul-99
GB_BA1:MSGY151	37036	AD000018	Mycobacterium tuberculosis sequence from clone y151.	Mycobacterium tuberculosis	36,634	10-DEC-1996
GB_BA1:ITCY130	32514	Z73902	Mycobacterium tuberculosis H37Rv complete genome; segment 59/162.	Mycobacterium tuberculosis	59,596	17-Jun-98
GB_BA1:MSGY151	37036	AD000018	Mycobacterium tuberculosis sequence from clone y151.	Mycobacterium tuberculosis	39,567	10-DEC-1996
GB_EST18:AA676822	524	AA676822	zf65c11.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:455156 3', mRNA sequence.	Homo sapiens	40,741	19-DEC-1997
GB_EST18:AA676822	524	AA676822	zf65c11.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:455156 3', mRNA sequence.	Homo sapiens	40,331	19-DEC-1997
GB_HTG4:AC010195	164935	AC010195	Homo sapiens chromosome 10 clone RPC111-587C2, *** SEQUENCING IN PROGRESS ***; 61 unordered pieces.	Homo sapiens	36,058	21-OCT-1999
GB_HTG4:AC010195	164935	AC010195	Homo sapiens chromosome 10 clone RPC111-587C2, *** SEQUENCING IN PROGRESS ***; 61 unordered pieces.	Homo sapiens	36,058	21-OCT-1999
GB_HTG1:CEY44A6	326074	Z98863	Caenorhabditis elegans chromosome V clone Y44A6, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	39,151	03-DEC-1998
GB_HTG1:CER08A5	51920	Z82281	Caenorhabditis elegans chromosome V clone R08A5, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	43,557	14-OCT-1998
GB_HTG1:CER08A5	51920	Z82281	Caenorhabditis elegans chromosome V clone R08A5, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	43,557	14-OCT-1998
GB_HTG1:CER08A5	51920	Z82281	Caenorhabditis elegans chromosome V clone R08A5, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	34,987	14-OCT-1998
GB_PL2:SPBC8D2	43757	AL022072	S. pombe chromosome II cosmid c8D2.	Schizosaccharomyces pombe	36,566	24-Nov-99
GB_PL1:AB004538	38911	AB004538	Schizosaccharomyces pombe 39 kb genomic DNA, clone c568.	Schizosaccharomyces pombe	36,148	15-Jul-97
GB_HTG6:AC009220	110000	AC009220	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS ***; 191 unordered pieces.	Homo sapiens	33,940	15-Sep-99
GB_GSS4:AQ694235	530	AQ694235	HS_5496_A1_D03_T7A RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=1072 Col=5 Row=G, genomic survey sequence.	Homo sapiens	37,259	6-Jul-99
GB_GSS11:AQ322059	519	AQ322059	RPC111-100G10.TV RPC1-11 Homo sapiens genomic clone RPC1-11-100G10, genomic survey sequence.	Homo sapiens	37,229	06-MAY-1999
GB_BA1:SCH24	41625	AL049626	Streptomyces coelicolor cosmid H24.	Streptomyces coelicolor	38,126	11-MAY-1999
GB_PR3:HS356B7	20733	AL031714	Human DNA sequence from clone 356B7 on chromosome 16. Contains the UBE21 gene for ubiquitin-conjugating enzyme E21 (homologous to yeast UBC9), and an RPS20 (40S Ribosomal protein S20) pseudogene. Contains ESTs, STSs, GSSs and a putative CpG island, complete sequence.	Homo sapiens	36,684	23-Nov-99
GB_PR3:HS356B7	20733	AL031714	Human DNA sequence from clone 356B7 on chromosome 16. Contains the UBE21 gene for ubiquitin-conjugating enzyme E21 (homologous to yeast UBC9), and an RPS20 (40S Ribosomal protein S20) pseudogene. Contains ESTs, STSs, GSSs and a putative CpG island, complete sequence.	Homo sapiens	39,621	23-Nov-99
GB_PR1:HSEF1A1	1815	X16870	Human DNA for elongation factor 1-alpha (clone lambda-1).	Homo sapiens	37,462	12-Jun-90
GB_PR3:HS25719	90695	AL031773	Human DNA sequence from clone 25719 on chromosome 6q25.1-26 Contains gene similar to Cytochrome B, CA repeat, GSS, complete sequence.	Homo sapiens	35,897	23-Nov-99
GB_PR3:HS28714	112831	Z95325	Human DNA sequence from clone 28714 on chromosome Xq21.1-21.33 Contains STSs and GSSs, complete sequence.	Homo sapiens	37,778	23-Nov-99

TABLE 4: ALIGNMENT RESULTS

GB_EST8:AA035251	493	AA035251	zk23d09.s1 Soares_pregnant_uterus_NbhPU Homo sapiens cDNA clone IMAGE:471377 3', mRNA sequence.	Homo sapiens	40,191	10-MAY-1997
GB_PR4:HUAC002394	107910	AC002394	Human Chromosome 16 BAC clone CIT987SK-A-211C6, complete sequence.	Homo sapiens	33,633	23-Nov-99
GB_OV:AF138905	1678	AF138905	Gallus gallus NK class homeodomain transcription factor NKX3.2 mRNA, complete cds.	Gallus gallus	30,556	4-Sep-99
GB_PAT:AR031772	30001	AR031772	Sequence 1 from patent US 5866410.	Unknown	34,304	29-Sep-99
GB_EST25:AI295169	507	AI295169	LP08720.3prime LP Drosophila melanogaster larval-early pupal pOT2 Drosophila melanogaster cDNA clone LP08720.3prime, mRNA sequence.	Drosophila melanogaster	36,279	01-DEC-1998
GB_EST25:AI297653	512	AI297653	LP12002.3prime LP Drosophila melanogaster larval-early pupal pOT2 Drosophila melanogaster cDNA clone LP12002.3prime, mRNA sequence.	Drosophila melanogaster	36,893	01-DEC-1998
GB_EST21:AA951454	677	AA951454	LD31920.5prime LD Drosophila melanogaster embryo pOT2 Drosophila melanogaster cDNA clone LD31920.5prime, mRNA sequence.	Drosophila melanogaster	44,167	24-Nov-98
GB_GSS12:AO365352	431	AO365352	nbxb0063L13f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0063L13f, genomic survey sequence.	Oryza sativa	34,577	3-Feb-99
GB_PR4:AC007544	119034	AC007544	Homo sapiens Human 12p11-37.2-54.4 BAC RPC111-12D15 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	Homo sapiens	34,350	29-OCT-1999
GB_PR4:AC007544	119034	AC007544	Homo sapiens Human 12p11-37.2-54.4 BAC RPC111-12D15 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	Homo sapiens	37,988	29-OCT-1999
GB_GSS1:AG009269	706	AG009269	Homo sapiens genomic DNA, 21q region, clone: 31C6X11, genomic survey sequence.	Homo sapiens	36,328	14-Apr-99
GB_GSS1:AG009269	706	AG009269	Homo sapiens genomic DNA, 21q region, clone: 31C6X11, genomic survey sequence.	Homo sapiens	37,391	14-Apr-99
GB_HTG2:AC007646	180133	AC007646	Drosophila melanogaster chromosome 3 clone BACR03J04 (D687) RPC1-98 03.J.4 map 87F-87F strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 10 unordered pieces.	Drosophila melanogaster	38,188	2-Aug-99
GB_HTG2:AC007646	180133	AC007646	Drosophila melanogaster chromosome 3 clone BACR03J04 (D687) RPC1-98 03.J.4 map 87F-87F strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 10 unordered pieces.	Drosophila melanogaster	38,188	2-Aug-99
GB_HTG4:AC009492	212394	AC009492	Homo sapiens clone NH0423F09. *** SEQUENCING IN PROGRESS ***. 2 unordered pieces.	Homo sapiens	38,762	29-OCT-1999
GB_EST7:W84105	361	W84105	T2969 MVAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma brucei rhodesiense cDNA 5', mRNA sequence.	Trypanosoma brucei rhodesiense	38,261	12-Aug-96
GB_EST7:W84105	361	W84105	T2969 MVAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma brucei rhodesiense cDNA 5', mRNA sequence.	Trypanosoma brucei rhodesiense	33,894	12-Aug-96
GB_RO:AC005240	41830	AC005240	Mus musculus clone UWGC:mapap from 14D1-D2 (T-cell Receptor Alpha Locus), complete sequence.	Mus musculus	41,463	8-Jul-98
GB_RO:AC004101	39491	AC004101	Mouse Cosmid ma53a016 from 14D1-D2, complete sequence.	Mus musculus	41,463	2-Feb-98
GB_HTG3:AC009837	162287	AC009837	Homo sapiens chromosome 17 clone 550_K_23 map 17. *** SEQUENCING IN PROGRESS ***. 13 unordered pieces.	Homo sapiens	36,131	2-Sep-99
GB_PR4:AC005919	156300	AC005919	Homo sapiens chromosome 17, clone hRPK.142_H_19, complete sequence.	Homo sapiens	39,822	18-Nov-98
GB_PR4:AC005919	156300	AC005919	Homo sapiens chromosome 17, clone hRPK.142_H_19, complete sequence.	Homo sapiens	38,319	18-Nov-98
GB_HTG1:CEY71A12_3110000	100793	AL021390	Caenorhabditis elegans chromosome I clone Y71A12. *** SEQUENCING IN PROGRESS ***. in unordered pieces.	Caenorhabditis elegans	36,884	15-Sep-99
GB_HTG4:AC010775	165565	AC010775	Homo sapiens clone 2_G_17. *** SEQUENCING IN PROGRESS ***. 14 unordered pieces.	Homo sapiens	38,670	20-OCT-1999
GB_HTG4:AC010775	165565	AC010775	Homo sapiens clone 2_G_17. *** SEQUENCING IN PROGRESS ***. 14 unordered pieces.	Homo sapiens	38,670	20-OCT-1999
GB_HTG4:AC010775	165565	AC010775	Homo sapiens clone 2_G_17. *** SEQUENCING IN PROGRESS ***. 14 unordered pieces.	Homo sapiens	36,131	20-OCT-1999
GB_PR3:AF070717	100793	AF070717	Homo sapiens BAC clone 393122 from 8q21, complete sequence.	Homo sapiens	38,826	2-Jul-98
GB_BA2:MPU34931	2571	U34931	Mycoplasma pulmonis FtsZ (ftsZ) gene, complete cds, methionyl-tRNA synthetase (metG) gene, partial cds.	Mycoplasma pulmonis	37,092	30-MAY-1996
GB_EST25:AI322057	638	AI322057	SWOv3MCAM09A04SK Onchocerca volvulus molting L3 larva cDNA (SL96MLW-Ovml.3)	Onchocerca volvulus	40,566	22-DEC-1998
GB_HTG3:AC010878	288945	AC010878	Onchocerca volvulus cDNA clone SWOv3MCAM09A04 5', mRNA sequence.	Onchocerca	36,496	25-Sep-99
GB_HTG3:AC010878	288945	AC010878	Homo sapiens clone NH0230E20. *** SEQUENCING IN PROGRESS ***. 65 unordered pieces.	Homo sapiens	36,496	25-Sep-99
GB_HTG3:AC010878	288945	AC010878	Homo sapiens clone NH0230E20. *** SEQUENCING IN PROGRESS ***. 65 unordered pieces.	Homo sapiens	36,496	25-Sep-99
GB_HTG3:AC010878	288945	AC010878	Homo sapiens clone NH0230E20. *** SEQUENCING IN PROGRESS ***. 65 unordered pieces.	Homo sapiens	36,354	25-Sep-99

TABLE 4: ALIGNMENT RESULTS

rx01271	1935	GB_BA2:U67549	14561	U67549	Methanococcus jannaschii	37,841	28-Jan-98
GB_PR2:HSJ836N17	111694	AL049539			Human DNA sequence from clone RP5-836N17 on chromosome 20q11.1-11.21, complete sequence.		
GB_OV:AF089743	30676	AF089743			Morone saxatilis homeodomain protein Hox-A7 (Hoxa7), homeodomain protein Hox-A5 (Hoxa5), and homeodomain protein Hox-A4 (Hoxa4) genes, complete cds.		
rx01273	1557	GB_PR3:AC004070	110192	AC004070	Human Chromosome X, complete sequence.	37,679	29-Jan-98
GB_PR2:AC002410	96217	AC002410			Human BAC clone RG264L19 from 7p15-p21, complete sequence.	38,587	11-Aug-97
GB_PR2:AC002410	96217	AC002410			Human BAC clone RG264L19 from 7p15-p21, complete sequence.	37,035	11-Aug-97
rx01275	1041	GB_EST8:C03322	357	C03322	C03322 Human heart cDNA (YNakamura) Homo sapiens cDNA clone 3NHC1263, mRNA sequence.	35,393	30-Jul-96
GB_EST4:R92079	410	R92079			yp96g02.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:195314 5', mRNA sequence.	37,010	25-Aug-95
GB_EST4:H57724	461	H57724			yr21a12.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:205918 5', mRNA sequence.	37,281	05-OCT-1995
rx01276	567	GB_OM:OCMYLC1	962	X54041	Rabbit mRNA for myosin light chain 1.		
GB_PR3:AC005383	123110	AC005383			Homo sapiens chromosome 10 clone C17987SK-1144G6 map 10q25.1, complete sequence.	40,503	26-Nov-90
GB_RO:RATPTPECA	2155	D78610			Rat mRNA for protein tyrosine phosphatase epsilon C, partial cds.	38,930	31-OCT-1998
rx01281	885	GB_HTG4:AC009366	199607	AC009366	Drosophila melanogaster chromosome 3L79D4 clone RPC198-48E10, *** SEQUENCING IN PROGRESS ***; 31 unordered pieces.	37,770	5-Feb-99
GB_HTG4:AC009366	199607	AC009366			Drosophila melanogaster chromosome 3L79D4 clone RPC198-48E10, *** SEQUENCING IN PROGRESS ***; 31 unordered pieces.	34,025	16-OCT-1999
GB_HTG4:AC009366	199607	AC009366			Drosophila melanogaster chromosome 3L79D4 clone RPC198-48E10, *** SEQUENCING IN PROGRESS ***; 31 unordered pieces.	34,386	16-OCT-1999
rx01282	903	GB_GSS11:AQ263970	363	AQ263970	CITB1-E1-2503H24, TF CITB1-E1 Homo sapiens genomic clone 2503H24, genomic survey sequence.	40,361	27-OCT-1998
GB_GSS1:FR0025959	603	AL018794			F.rubripes GSS sequence, clone 165E10aE1, genomic survey sequence.		
GB_GSS9:AQ102435	334	AQ102435			HS_3038_B2_D09_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-3038 Col=18 Row=H, genomic survey sequence.	34,824	10-DEC-1997
rx01294	789	GB_PL2:AC007258	144422	AC007258	Arabidopsis thaliana chromosome I BAC F23H11 genomic sequence, complete sequence.	39,453	16-Jun-99
GB_EST17:T04634	491	T04634			681 Lambda-PRL1 Arabidopsis thaliana cDNA clone SBD1T7P, mRNA sequence.	38,693	6-Nov-97
GB_PL2:AC007258	144422	AC007258			Arabidopsis thaliana chromosome I BAC F23H11 genomic sequence, complete sequence.	36,170	16-Jun-99
rx01295	1098	GB_HTG3:AC009803	235360	AC009803	Homo sapiens clone RPC111-1028N23, *** SEQUENCING IN PROGRESS ***; 47 unordered pieces.	38,662	24-Sep-99
GB_HTG3:AC009803	235360	AC009803			Homo sapiens clone RPC111-1028N23, *** SEQUENCING IN PROGRESS ***; 47 unordered pieces.	38,662	24-Sep-99
rx01296	1206	GB_IN2:AC004333	63178	AC004333	Drosophila melanogaster DNA sequence (P1 DS05969 (D229)), complete sequence.	39,252	20-Jun-98
GB_PR4:AC004961	68130	AC004961			Homo sapiens clone DJ1098J04, complete sequence.	39,409	5-Jun-99
GB_GSS14:AQ525299	447	AQ525299			HS_5227_B2_H03_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=803 Col=6 Row=P, genomic survey sequence.	37,330	11-MAY-1999
rx01301	648	GB_BA2:BFU78108	1900	U78108	Bacteroides fragilis NAD(H) glutamate dehydrogenase (gdhB) gene, complete cds.	40,509	15-Jul-98
GB_PR4:AC007051	167810	AC007051			Homo sapiens chromosome 3, clone hRPK.44_A.1, complete sequence.	38,498	11-Jun-99
GB_PR4:AC007919	184989	AC007919			Homo sapiens 3q26.2-27 BAC RPC111-408H1 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	35,139	09-OCT-1999
GB_EST1:AA252547	454	AA252547			zpf7g09.r1 Striatagene HeLa cell s3 937216 Homo sapiens cDNA clone IMAGE:827232 5' similar to gb.D13748 EUKARYOTIC INITIATION FACTOR 4A-1 (HUMAN), mRNA sequence.	38,538	12-MAR-1997
rx01304	609	GB_HTG2:AC006779	119562	AC006779	Caenorhabditis elegans clone Y47D7, *** SEQUENCING IN PROGRESS ***; 32 unordered pieces.	32,888	25-Feb-99
GB_HTG2:AC006779	119562	AC006779			Caenorhabditis elegans clone Y47D7, *** SEQUENCING IN PROGRESS ***; 32 unordered pieces.	32,888	25-Feb-99

## TABLE 4: ALIGNMENT RESULTS

TABLE 4. ALIGNMENT RESULTS									
rx01306	1131	GB_VI:BHT1UL GB_GSS8:AQ047475 GB_PL1:CR1433P	37000 420 1464	Z78205 AQ047475 X79445	Bovine herpesvirus type 1 UL22-35 genes. cLM-1c5-t cLM Giardia intestinalis genomic, genomic survey sequence. C.reinhardtii mRNA for 14-3-3 protein.	Bovine herpesvirus 1 Giardia intestinalis Chlamydomonas reinhardtii	38,047 38,902 39,640	14-Aug-96 14-Jul-98 06-OCT-1995	
rx01310	450	GB_PR3:HSJ878113 GB_HTG3:AC010530 GB_HTG3:AC010530 GB_HTG3:AC010530 GB_HTG3:AC010530	122400 274508 274508 274508 274508	AL049591 AC010530 AC010530 AC010530 AC010530	Human DNA sequence from clone 878113 on chromosome Xq23-25 Contains a pseudogene similar to alpha tubulin, ESTs, STSs, GSSs, complete sequence. Homo sapiens chromosome 16 clone RPC1-11_167P11, *** SEQUENCING IN PROGRESS ***, 19 unordered pieces. Homo sapiens chromosome 16 clone RPC1-11_167P11, *** SEQUENCING IN PROGRESS ***, 19 unordered pieces. Homo sapiens chromosome 16 clone RPC1-11_167P11, *** SEQUENCING IN PROGRESS ***, 19 unordered pieces.	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	35,624 38,462 38,462 38,462 32,063	23-Nov-99 15-Sep-99 15-Sep-99 15-Sep-99 15-Sep-99	
rx01315	774	GB_BA2:AF031037 GB_GSS11:AQ326599 GB_HTG1:PFMAL13PA	1472 662 80518	AF031037 AQ326599 AL109815	Neisseria meningitidis chloramphenicol acetyltransferase gene, complete cds. nbxb0037P01r CUGI Rice BAC Library Oryza sativa genomic clone nbxb0037P01r, genomic survey sequence. Plasmodium falciparum chromosome 13 strain 3D7, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Neisseria meningitidis Oryza sativa Plasmodium falciparum	35,638 34,817 34,935	21-Apr-98 8-Jan-99 19-Aug-99	
rx01316	426	GB_HTG3:AC010133 GB_HTG3:AC010133 GB_HTG3:AC010133 GB_HTG3:AC010133 GB_BA2:U67560	154773 154773 154773 12215 12215	AC010133 AC010133 AC010133 U67560 U67560	Homo sapiens clone NH0118E09, *** SEQUENCING IN PROGRESS ***, 5 unordered pieces. Homo sapiens clone NH0118E09, *** SEQUENCING IN PROGRESS ***, 5 unordered pieces. Homo sapiens clone NH0118E09, *** SEQUENCING IN PROGRESS ***, 5 unordered pieces. Methanococcus jannaschii section 102 of 150 of the complete genome.	Homo sapiens Homo sapiens Homo sapiens Methanococcus jannaschii	33,414 33,414 38,186 38,476	13-Sep-99 13-Sep-99 13-Sep-99 28-Jan-98	
rx01318	1425	GB_BA1:MTV038 GB_IN1:CELC41A3 GB_HTG3:AC011298 GB_BA1:PAL249201 GB_BA1:PAL249201	16094 37149 205637 3390 3390	AL021933 U41541 AC011298 AJ249201 AJ249201	Methanococcus jannaschii section 102 of 150 of the complete genome. Mycobacterium tuberculosis H37Rv complete genome; segment 24/162. Caenorhabditis elegans cosmid C41A3. Homo sapiens clone NH0118M12, *** SEQUENCING IN PROGRESS ***, 19 unordered pieces. Prevotella albensis ftsQ (partial), ftsA and ftsZ genes and ORF-fts (partial). Prevotella albensis ftsQ (partial), ftsA and ftsZ genes and ORF-fts (partial).	Methanococcus jannaschii Mycobacterium tuberculosis Caenorhabditis elegans Homo sapiens Prevotella albensis Prevotella albensis	35,867 51,266 36,887 35,760 37,708 37,474	28-Jan-98 17-Jun-98 08-DEC-1995 05-OCT-1999 27-Aug-99	
rx01331	2520	GB_HTG1:HSJ02553 GB_HTG1:HSJ02553 GB_PR3:AC005304 GB_BA1:CGU43536	142807 142807 164017 3464	AJ002553 AJ002553 AC005304 U43536	Homo sapiens chromosome 11 clone 1118122 map q13, *** SEQUENCING IN PROGRESS ***, in unordered pieces. Homo sapiens chromosome 11 clone 1118122 map q13, *** SEQUENCING IN PROGRESS ***, in unordered pieces. Homo sapiens chromosome 17, clone hRPK_131_K_5, complete sequence. Corynebacterium glutamicum heat shock, ATP-binding protein (clpB) gene, complete cds.	Homo sapiens Homo sapiens Homo sapiens Corynebacterium glutamicum	26,237 26,237 35,318 38,075	14-Nov-97 14-Nov-97 25-Jul-98 13-MAR-1997	
rx01336	726	GB_OM:BTPLRP GB_OM:BTPLRP GB_BA1:AP000064 GB_BA1:AP000064 GB_EST1:T05458	924 924 247695 247695 363	X15975 X15975 AP000064 AP000064 T05458	Bovine mRNA for prolactin (PRL) related protein. Bovine mRNA for prolactin (PRL) related protein. Aeropyrum pernix genomic DNA, section 77. Aeropyrum pernix genomic DNA, section 77. EST03347 Fetal brain, Striatagene (cat#936206) Homo sapiens cDNA clone HFBY86, mRNA sequence.	Bos taurus Bos taurus Aeropyrum pernix Aeropyrum pernix Homo sapiens	40,901 39,580 37,881 36,161 43,798	22-DEC-1994 22-DEC-1994 22-Jun-99 22-Jun-99 30-Jun-93	



TABLE 4: ALIGNMENT RESULTS

rx01337	1536	GB_PR2:AP000215	100000	AP000215	Homo sapiens genomic DNA, chromosome 21q22.3-ter, Ter region, clone f27E1-T1136, segment 1/4, complete sequence.	Homo sapiens	36,438	20-Nov-99
		GB_PR2:AP000215	100000	AP000215	Homo sapiens genomic DNA, chromosome 21q22.3-ter, Ter region, clone f27E1-T1136, segment 1/4, complete sequence.	Homo sapiens	35,827	20-Nov-99
		GB_PR2:AP000337	53553	AP000337	Homo sapiens genomic DNA, chromosome 21q22.3-ter, Ter region, clone T1957, complete sequence.	Homo sapiens	36,043	20-Nov-99
rx01342	626	GB_EST35:A1814229	544	A1814229	wj70e01.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2408184.3', mRNA sequence.	Homo sapiens	36,803	24-Aug-99
		GB_EST35:A1814229	544	A1814229	wj70e01.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2408184.3', mRNA sequence.	Homo sapiens	36,688	24-Aug-99
rx01348	615	GB_HTG3:AC011246	210407	AC011246	Homo sapiens clone NH0498020, *** SEQUENCING IN PROGRESS ***; 29 unordered pieces.	Homo sapiens	39,101	04-OCT-1999
		GB_HTG3:AC011246	210407	AC011246	Homo sapiens clone NH0498020, *** SEQUENCING IN PROGRESS ***; 29 unordered pieces.	Homo sapiens	39,101	04-OCT-1999
		GB_PL1:ZMB32120	3093	X54212	Z. mays mRNA for b-32 protein, putative regulatory factor of zein expression (clone b-32.120).	Zea mays	35,902	18-Jan-93
rx01349	900	GB_IN1:CELK10G6	33588	AF016669	Caenorhabditis elegans cosmid K10G6.	Caenorhabditis elegans	40,251	8-Aug-97
		GB_IN1:CELK10G6	33588	AF016669	Caenorhabditis elegans cosmid K10G6.	Caenorhabditis elegans	36,842	8-Aug-97
rx01357	426	GB_BA1:MTCY274	39991	Z74024	Mycobacterium tuberculosis H37Rv complete genome; segment 126/162.	Mycobacterium tuberculosis	36,058	19-Jun-98
		GB_BA1:MLCB250	40603	Z97369	Mycobacterium leprae cosmid B250.	Mycobacterium leprae	59,294	27-Aug-99
		GB_BA2:SKZ86111	7860	Z86111	Streptomyces lividans rpsP, trmD, rplS, sipW, sipX, sipY, sipZ, mult genes and 4 open reading frames.	Streptomyces lividans	57,882	27-OCT-1999
rx01359	1305	GB_BA1:D87820	7217	D87820	Vibrio cholerae non-O1 gene for N-acetylglucosamine 6-phosphate deacetylase, NagC, NagE, complete cds.	Vibrio cholerae non-O1	38,152	17-OCT-1997
		GB_PR3:AC004540	131757	AC004540	Homo sapiens PAC clone DJ1066K24 from 7p15, complete sequence.	Homo sapiens	35,714	9-Apr-98
		GB_PR2:HSU44119	450	U44119	Human der(9) chromosome breakpoint region: alpha 1(V) collagen chain (COL5A1) gene, partial intron 24, and imperfect LINE-1 element of Xp21.2.	Homo sapiens	38,085	14-Jan-97
rx01362	3677	GB_EST30:A1665031	602	A1665031	605005H05.x1 605 - Endospore cDNA library from Schmidt lab Zea mays cDNA, mRNA sequence.	Zea mays	42,308	11-MAY-1999
		GB_EST34:A1795319	661	A1795319	605005H05.y2 605 - Endospore cDNA library from Schmidt lab Zea mays cDNA, mRNA sequence.	Zea mays	38,729	2-Jul-99
		GB_HTG3:AC008387	151804	AC008387	Homo sapiens chromosome 5 clone CIT-HSPC_229p9, *** SEQUENCING IN PROGRESS ***; 74 unordered pieces.	Homo sapiens	38,023	3-Aug-99
rx01366	513	GB_IN1:AF007166	1392	AF007166	Anopheles gambiae serine protease 14D mRNA, complete cds.	Anopheles gambiae	41,379	14-Jul-97
		GB_GSS13:AQ475498	476	AQ475498	CITBI-E1-2589F13.TR CITBI-E1 Homo sapiens genomic clone 2589F13, genomic survey sequence.	Homo sapiens	40,511	23-Apr-99
rx01367	525	GB_IN1:AF007166	1392	AF007166	Anopheles gambiae serine protease 14D mRNA, complete cds.	Anopheles gambiae	40,000	14-Jul-97
		GB_EST28:A1489958	591	A1489958	EST248297 tomato ovary, TAMU Lycopersicon esculentum cDNA clone cLED14O8, mRNA sequence.	Lycopersicon esculentum	38,186	29-Jun-99
		GB_GSS13:AQ475498	476	AQ475498	CITBI-E1-2589F13.TR CITBI-E1 Homo sapiens genomic clone 2589F13, genomic survey sequence.	Homo sapiens	37,405	23-Apr-99
		GB_EST28:A1489958	591	A1489958	EST248297 tomato ovary, TAMU Lycopersicon esculentum cDNA clone cLED14O8, mRNA sequence.	Lycopersicon esculentum	35,882	29-Jun-99
rx01370								
rx01372	614	GB_HTG5:AC011644	164746	AC011644	Homo sapiens clone 14_K_21 *** SEQUENCING IN PROGRESS ***; 8 unordered pieces.	Homo sapiens	33,715	5-Nov-99
		GB_HTG5:AC011644	164746	AC011644	Homo sapiens clone 14_K_21 *** SEQUENCING IN PROGRESS ***; 8 unordered pieces.	Homo sapiens	35,821	5-Nov-99
rx01378	1878	GB_IN1:DMBR42117	155168	AL121806	Drosophila melanogaster clone BACR42117.	Drosophila melanogaster	36,373	10-OCT-1999
		GB_IN2:AC005714	177740	AC005714	Drosophila melanogaster, chromosome 2R, region 58D4-58E2, BAC clone BACR48M13, complete sequence.	Drosophila melanogaster	36,234	01-MAY-1999

TABLE 4: ALIGNMENT RESULTS

rx01379	1042	GB_IN2:AC005639	188288	AC005639	Drosophila melanogaster, chromosome 2R, region 59E3-59F4, BAC clone BACR48M01, complete sequence.	Drosophila melanogaster	36,275	6-Jan-99
rx01380	2322	GB_HTG4:AC010031	132106	AC010031	Drosophila melanogaster chromosome 3L70C1 clone RPC198-2M20, *** SEQUENCING IN PROGRESS ***; 69 unordered pieces.	Drosophila melanogaster	35,368	16-OCT-1999
rx01384	1839	GB_HTG4:AC010031	132106	AC010031	Drosophila melanogaster chromosome 3L70C1 clone RPC198-2M20, *** SEQUENCING IN PROGRESS ***; 69 unordered pieces.	Drosophila melanogaster	35,368	16-OCT-1999
		GB_HTG3:AC009932	68745	AC009932	Homo sapiens clone 114_O_12, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	36,337	05-OCT-1999
		GB_BA1:SC5F2A	40105	AL049587	Streptomyces coelicolor cosmid 5F2A.	Streptomyces coelicolor	39,699	24-MAY-1999
		GB_BA1:BPTEX	2701	X95386	B pertussis tex gene.	Bordetella pertussis	59,687	31-Jul-96
		GB_BA1:SC5F2A	40105	AL049587	Streptomyces coelicolor cosmid 5F2A.	Streptomyces coelicolor	41,024	24-MAY-1999
		GB_GSS15:AQ613816	598	AQ613816	HS_5118_B1_F06_T7A RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=694 Col=11 Row=L, genomic survey sequence.	Homo sapiens	39,300	15-Jun-99
		GB_GSS15:AQ613816	598	AQ613816	HS_5118_B1_F06_T7A RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=694 Col=11 Row=L, genomic survey sequence.	Homo sapiens	36,455	15-Jun-99
rx01390	780	GB_BA2:CGI012293	2952	AJ012293	Corynebacterium glutamicum livD gene.	Corynebacterium glutamicum	40,413	01-OCT-1999
		GB_PR4:AC006213	160754	AC006213	Homo sapiens, clone hRPK.15_A_1, complete sequence.	Homo sapiens	35,724	16-Jan-99
		GB_PL2:ATF24G24	99856	AL049488	Arabidopsis thaliana DNA chromosome 4, BAC clone F24G24 (ESSA project).	Arabidopsis thaliana	35,925	27-Aug-99
rx01391	813	GB_HTG2:AC006171	44733	AC006171	Homo sapiens chromosome 10 clone LA10NC01_15_E_11 map 10q26.3, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Homo sapiens	35,031	09-DEC-1998
		GB_HTG2:AC006171	44733	AC006171	Homo sapiens chromosome 10 clone LA10NC01_15_E_11 map 10q26.3, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Homo sapiens	35,031	09-DEC-1998
		GB_HTG2:AC006171	44733	AC006171	Homo sapiens chromosome 10 clone LA10NC01_15_E_11 map 10q26.3, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Homo sapiens	38,035	09-DEC-1998
rx01396	381	GB_BA2:AE000775	14358	AE000775	Aquifex aeolicus section 107 of 109 of the complete genome.	Aquifex aeolicus	43,085	25-MAR-1998
		GB_EST26:A1329024	471	A1329024	agc02ne.f11 Neurospora crassa evening cDNA library Neurospora crassa cDNA clone agc02ne 5' mRNA sequence.	Neurospora crassa	39,286	28-DEC-1998
		GB_EST26:A1329043	401	A1329043	agc06ne.f11 Neurospora crassa evening cDNA library Neurospora crassa cDNA clone agc06ne 5' mRNA sequence.	Neurospora crassa	44,484	28-DEC-1998
rx01397	543	GB_BA2:AE000775	14358	AE000775	Aquifex aeolicus section 107 of 109 of the complete genome.	Aquifex aeolicus	41,121	25-MAR-1998
		GB_BA1:BSTMSPRS	3211	X16518	B subtilis prs.tms and ctc (partial) genes for PRPP synthetase and two undefined gene products.	Bacillus subtilis	41,255	30-Sep-93
		GB_EST4:H21691	382	H21691	y129b01.r1 Soares breast 3NbrHst1 Homo sapiens cDNA clone IMAGE:159625 5' similar to SP-XPEC_CERAE P33194 POSSIBLE DNA-REPAIR PROTEIN XP-E.; mRNA sequence.	Homo sapiens	35,407	6-Jul-95
rx01400	1623	GB_GSS11:AQ264714	353	AQ264714	CITBI-E1-2502C17 TR CITBI-E1 Homo sapiens genomic clone 2502C17, genomic survey sequence.	Homo sapiens	40,227	27-OCT-1998
rx01401	384	GB_BA1:TFENTRA	2803	M59480	Thiobacillus ferrooxidans nitrogen metabolism regulator (nitrA) gene, complete cds.	Thiobacillus ferrooxidans	39,960	26-Apr-93
		GB_BA1:NPSDHCDBA	5596	Y07709	N pharaonis sdhC, sdhD, sdhB and sdhA genes.	Nastronomas pharaonis	38,003	2-Sep-96
		GB_HTG4:AC006583	110000	AC006583	Homo sapiens chromosome 3p21.3 clone RPC111-491D6, *** SEQUENCING IN PROGRESS ***; 90 unordered pieces.	Homo sapiens	32,718	21-OCT-1999
		GB_HTG4:AC006583	110000	AC006583	Homo sapiens chromosome 3p21.3 clone RPC111-491D6, *** SEQUENCING IN PROGRESS ***; 90 unordered pieces.	Homo sapiens	32,718	21-OCT-1999
		GB_HTG4:AC011291	173585	AC011291	Homo sapiens chromosome unknown clone NH0067G07, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	33,596	29-OCT-1999
rx01402	597	GB_BA1:CAJ10319	5368	AJ010319	Corynebacterium glutamicum amfP, glnB, glnD genes and partial flsY and srp genes.	Corynebacterium glutamicum	39,389	14-MAY-1999
		GB_HTG4:AC010066	187240	AC010066	Drosophila melanogaster chromosome 3L72A4 clone RPC198-25O1, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Drosophila melanogaster	33,681	16-OCT-1999
		GB_HTG4:AC010066	187240	AC010066	Drosophila melanogaster chromosome 3L72A4 clone RPC198-25O1, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Drosophila melanogaster	33,681	16-OCT-1999

TABLE 4: ALIGNMENT RESULTS

rx01403	771	GB_BA1:SC6E10	23990	AL109661	Streptomyces coelicolor cosmid 6E10.	Streptomyces coelicolor A3(2)	39,136	5-Aug-99
		GB_GSS12:AQ396728	608	AQ396728	mgxb0002E02f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0002E02f, genomic survey sequence.	Magnaporthe grisea	36,626	06-MAR-1999
		GB_BA1:SC6E10	23990	AL109661	Streptomyces coelicolor cosmid 6E10.	Streptomyces coelicolor A3(2)	41,403	5-Aug-99
rx01405	579	GB_PL2:AF111709	52684	AF111709	Oryza sativa subsp. indica Retrosat 1 retrotransposon and Ty3-Gypsy type Retrosat 2 retrotransposon, complete sequences; and unknown genes.	Oryza sativa subsp. indica	34,888	26-Apr-99
		GB_GSS3:B88760	696	B88760	RPC111-24L19.TPC RPC111 Homo sapiens genomic clone RPC111-24L19, genomic survey sequence.	Homo sapiens	37,204	9-Apr-99
rx01409	845	GB_STS:G52436	696	G52436	SHGC-85004 Human Homo sapiens STS genomic, sequence tagged site.	Homo sapiens	37,204	25-Jun-99
		GB_HTG5:AC008019	190459	AC008019	Mus musculus, *** SEQUENCING IN PROGRESS ***; 16 unordered pieces.	Mus musculus	39,924	16-Nov-99
		GB_PR4:AC006236	127593	AC006236	Homo sapiens chromosome 17, clone hCIT.162.E.12, complete sequence.	Homo sapiens	33,907	29-DEC-1998
		GB_HTG5:AC008019	190459	AC008019	Mus musculus, *** SEQUENCING IN PROGRESS ***; 16 unordered pieces.	Mus musculus	40,819	16-Nov-99
rx01413	723	GB_OV:AF038947	1817	AF038947	Ambystoma tigrinum red cone visual pigment mRNA, complete cds.	Ambystoma tigrinum	35,846	16-DEC-1998
		GB_BA1:MSGB577COS 37770	10264	L01263	M. leprae genomic dna sequence, cosmid b577.	Mycobacterium leprae	38,042	14-Jun-96
		GB_PR3:AC002558	102064	AC002558	Homo sapiens chromosome 17, clone hRPC867C24, complete sequence.	Homo sapiens	36,820	31-OCT-1997
rx01414	630	GB_PL1:SC5610	37730	Z38060	S.cerevisiae chromosome IX sequence derived from lambda clones 5610-5004.	Saccharomyces cerevisiae	38,782	20-Aug-97
		GB_PL1:YSCDIN1	2969	M58012	S.cerevisiae ribonucleotide reductase DNA damage-inducible regulatory subunit (DIN1) gene, 5' end.	Saccharomyces cerevisiae	35,877	27-Apr-93
		GB_EST18:AA689161	624	AA689161	ET2100 Trypanosoma brucei rhodesiense ZAP II library Trypanosoma brucei rhodesiense cDNA 5', mRNA sequence.	Trypanosoma brucei rhodesiense	37,150	15-DEC-1997
rx01417	720	GB_PL1:SC8419	30507	Z49701	S.cerevisiae chromosome IV cosmid 8419.	Saccharomyces cerevisiae	36,248	11-Aug-97
		GB_PL1:SCPRP28G	2452	X56934	S.cerevisiae PRP28 gene.	Saccharomyces cerevisiae	45,036	9-Feb-95
rx01421		GB_BA1:BACPEPFA	4654	M29035	B.subtilis bacillopeptidase F (bpf) gene, complete cds.	Bacillus subtilis	37,324	06-MAR-1995
rx01422	1095	GB_EST24:AU033392	344	AU033392	AU033392 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium discoideum cDNA clone SLA715, mRNA sequence.	Dictyostelium discoideum	38,953	28-Apr-99
		GB_EST14:AA399243	301	AA399243	z157d02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:728435 3' similar to SW:NIDO_HUMAN P14543 NIDOGN PRECURSOR., mRNA sequence.	Homo sapiens	45,183	12-Aug-97
rx01425	1008	GB_PR2:AC002037	53897	AC002037	Human Chromosome 11 Overlapping Cosmids cSRL72g7 and cSRL140b8, complete sequence.	Homo sapiens	37,500	6-Aug-97
		GB_PR4:AC005908	196501	AC005908	Homo sapiens 12p13.3 BAC RPC111-476M19 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	Homo sapiens	34,410	20-Jan-99
		GB_PR4:AF111169	183916	AF111169	Homo sapiens chromosome 14 BAC containing gene for KIAA0759 and other possible new transcripts, complete sequence.	Homo sapiens	40,635	19-Jul-99
		GB_HTG2:AC008284	146797	AC008284	Drosophila melanogaster chromosome 3 clone BACR03M22 (D1000) RPC1-98 03.M.22 map 96C-96D strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 137 unordered pieces.	Drosophila melanogaster	35,822	2-Aug-99
rx01429	741	GB_EST25:AI300084	635	AI300084	qh59f02.x1 NCL.CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1902555 3', mRNA sequence.	Homo sapiens	37,753	1-Feb-99
		GB_EST18:AA706612	949	AA706612	ah26c02.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone 1239938 3', mRNA sequence.	Homo sapiens	37,361	12-Jan-99
rx01434	3075	GB_EST35:AI817084	598	AI817084	wj76g07.x1 NCL.CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2408796 3', mRNA sequence.	Homo sapiens	38,627	24-Aug-99
		GB_PL1:CAC49C10	37825	AL033497	C.albicans cosmid Ca49C10.	Candida albicans	34,724	10-Nov-98
		GB_GSS3:BI0423	1217	BI0423	F19F22-T7 IGF Arabidopsis thaliana genomic clone F19F22, genomic survey sequence.	Arabidopsis thaliana	36,003	14-MAY-1997
		GB_PL1:CAC49C10	37825	AL033497	C.albicans cosmid Ca49C10.	Candida albicans	33,794	10-Nov-98

TABLE 4: ALIGNMENT RESULTS

rx01439	669	GB_BA2:AF049897	9196	AF049897	Corynebacterium glutamicum N-acetylglutamate kinase (argB), ornithine acetyltransferase (argJ), N-acetylglutamate kinase (argB), acetylornithine transaminase (argD), ornithine carbamoyltransferase (argF), arginine repressor (argR), argininosuccinate synthase (argG), and argininosuccinate lyase (argH) genes, complete cds.	Corynebacterium glutamicum	41,564	1-Jul-98
		GB_BA1:CGARGCJBD	4355	X86157	C. glutamicum argC, argJ, argB, argD, and argF genes.	Corynebacterium glutamicum	41,564	25-Jul-96
		GB_BA2:AF049897	9196	AF049897	Corynebacterium glutamicum N-acetylglutamate kinase (argB), ornithine acetyltransferase (argJ), N-acetylglutamate kinase (argB), acetylornithine transaminase (argD), ornithine carbamoyltransferase (argF), arginine repressor (argR), argininosuccinate synthase (argG), and argininosuccinate lyase (argH) genes, complete cds.	Corynebacterium glutamicum	39,486	1-Jul-98
rx01440	1131	GB_PAT:E16763	2517	E16763	gDNA encoding aspartate transferase (AAT).	Corynebacterium glutamicum	42,188	28-Jul-99
		GB_EST20:AA870569	423	AA870569	vq23c09.r1 Barstead stromal cell line MPLRB8 Mus musculus cDNA clone IMAGE:1095088 5' similar to TR:Q14468 Q14468 KIAA0047 ; mRNA sequence.	Mus musculus	42,066	16-MAR-1998
		GB_EST18:AA726390	436	AA726390	vu40b04.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone IMAGE:1193839 5' mRNA sequence.	Mus musculus	41,228	2-Jan-98
rx01441	1152	GB_EST32:AV053763	252	AV053763	AV053763 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA clone 1810028E06, mRNA sequence.	Mus musculus	45,238	23-Jun-99
		GB_EST20:AA870569	423	AA870569	vq23c09.r1 Barstead stromal cell line MPLRB8 Mus musculus cDNA clone IMAGE:1095088 5' similar to TR:Q14468 Q14468 KIAA0047 ; mRNA sequence.	Mus musculus	39,643	16-MAR-1998
		GB_EST18:AA726390	436	AA726390	vu40b04.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone IMAGE:1193839 5' mRNA sequence.	Mus musculus	39,953	2-Jan-98
rx01445	1116	GB_PR4:AC007786	225061	AC007786	Homo sapiens chromosome 19, BAC 41855 (CIT-B-3204), complete sequence.	Homo sapiens	39,545	11-Jun-99
		GB_PR3:AC003957	126581	AC003957	Homo sapiens chromosome 17, clone hCIT.457_L_16, complete sequence.	Homo sapiens	36,579	09-MAY-1998
		GB_OV:AF108420	46626	AF108420	Fugu rubripes prohormone convertase PACE4 (PACE4) gene, partial cds; and 1-aminocyclopropane-carboxylate synthase (ACC), recombination-activating protein 1 (RAG1), and recombination-activating protein 2 (RAG2) genes, complete cds.	Fugu rubripes	39,187	17-MAR-1999
rx01447	972	GB_PL2:AF049112	1436	AF049112	Zea mays retrotransposon Cinfu1 prpol mRNA, partial cds.	Zea mays	38,268	1-Feb-99
		GB_EST36:AV190070	360	AV190070	AV190070 Yuji Kohara unpublished cDNA: Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk566e10 5', mRNA sequence.	Caenorhabditis elegans	40,390	22-Jul-99
		GB_EST9:AA103228	477	AA103228	mo22h07.r1 Life Tech mouse embryo 13 5dpc 10666014 Mus musculus cDNA clone IMAGE:554365 5' similar to gb:J00068 ACTIN, ALPHA SKELETAL MUSCLE (HUMAN); gb:M12866 Mouse skeletal muscle actin mRNA, complete cds (MOUSE); mRNA sequence.	Mus musculus	40,411	29-OCT-1996
rx01448	1290	GB_BA1:CGFDA	3371	X17313	Corynebacterium glutamicum tda gene for fructose-bisphosphate aldolase (EC 4.1.2.13).	Corynebacterium glutamicum	100,000	12-Sep-93
		GB_PL2:ENU75428	4443	U75428	Emericella nidulans acid trehalase precursor (treA) gene, complete cds.	Emericella nidulans	36,271	15-MAY-1997
		GB_BA1:CGFDA	3371	X17313	Corynebacterium glutamicum tda gene for fructose-bisphosphate aldolase (EC 4.1.2.13).	Corynebacterium glutamicum	37,872	12-Sep-93
rx01452	402	GB_BA1:MTV017	67200	AL021897	Mycobacterium tuberculosis H37Rv complete genome, segment 48/162.	Mycobacterium tuberculosis	57,246	24-Jun-99
		GB_BA1:MLCB1222	34714	AL049491	Mycobacterium leprae cosmid B1222.	Mycobacterium leprae	34,872	27-Aug-99
		GB_BA2:S71532	914	S71532	che=cholesterol esterase [Streptomyces lavendulae, H646-SY2, Genomic, 914 nt].	Streptomyces lavendulae	38,701	8-Nov-94
rx01456	645	GB_GSS3:B16150	663	B16150	347A15.TP C17978SKA1 Homo sapiens genomic clone A-347A15, genomic survey sequence.	Homo sapiens	34,394	4-Jun-98
		GB_HTG3:AC004157	132090	AC004157	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS *** 9 unordered pieces.	Plasmodium falciparum	34,165	23-Sep-99
		GB_HTG3:AC004157	132090	AC004157	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS *** 9 unordered pieces.	Plasmodium falciparum	34,165	23-Sep-99
rx01457	798	GB_HTG3:AC011009	158335	AC011009	Homo sapiens clone 2_L_22, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	39,490	29-Sep-99
		GB_HTG3:AC011009	158335	AC011009	Homo sapiens clone 2_L_22, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	39,490	29-Sep-99
		GB_HTG3:AC011009	158335	AC011009	Homo sapiens clone 2_L_22, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	39,125	29-Sep-99

**TABLE 4: ALIGNMENT RESULTS**

rx01459	933	GB_HTG3:AC008670	113564	AC008670	Homo sapiens chromosome 5 clone CIT978SKB_3601, *** SEQUENCING IN PROGRESS ***; 43 unordered pieces.	Homo sapiens	39,022	3-Aug-99
		GB_HTG3:AC008670	113564	AC008670	Homo sapiens chromosome 5 clone CIT978SKB_3601, *** SEQUENCING IN PROGRESS ***; 43 unordered pieces.	Homo sapiens	39,022	3-Aug-99
		GB_IN2:CELLT05A8	33896	AF040652	Caenorhabditis elegans cosmid T05A8.	Caenorhabditis elegans	35,699	16-Jun-99
rx01460	417	GB_HTG3:AC009186	48600	AC009186	Homo sapiens chromosome 5 clone CIT978SKB_148114, *** SEQUENCING IN PROGRESS ***; 3 ordered pieces.	Homo sapiens	41,278	07-OCT-1999
		GB_HTG3:AC009186	48600	AC009186	Homo sapiens chromosome 5 clone CIT978SKB_148114, *** SEQUENCING IN PROGRESS ***; 3 ordered pieces.	Homo sapiens	41,278	07-OCT-1999
		GB_HTG3:AC009186	48600	AC009186	Homo sapiens chromosome 5 clone CIT978SKB_148114, *** SEQUENCING IN PROGRESS ***; 3 ordered pieces.	Homo sapiens	35,162	07-OCT-1999
		GB_HTG3:AC009186	48600	AC009186	Homo sapiens chromosome 5 clone CIT978SKB_148114, *** SEQUENCING IN PROGRESS ***; 3 ordered pieces.	Homo sapiens	37,956	23-Nov-99
rx01463	1287	GB_PR3:HS326L13	127247	Z82170	Human DNA sequence from PAC 326L13 containing brain-4 mRNA ESTs and polymorphic CA repeat.	Homo sapiens	38,035	23-Nov-99
		GB_HTG1:HS439A6	49379	AL031723	Homo sapiens chromosome 16 clone LA16-439A6, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,035	23-Nov-99
		GB_HTG1:HS439A6	49379	AL031723	Homo sapiens chromosome 16 clone LA16-439A6, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	39,806	3-Jun-99
rx01469	1155	GB_EST31:A1701691	349	A1701691	we81c04 x1 Soares_NFL_T_OBC_S1 Homo sapiens cDNA clone IMAGE:2347494 3' similar to gb:L19886_ma1 MACROPHAGE MIGRATION INHIBITORY FACTOR (HUMAN); mRNA sequence.	Homo sapiens	42,705	14-Aug-97
		GB_EST15:AA480256	389	AA480256	ne31f04.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:898975 3' similar to gb:L19886_ma1 MACROPHAGE MIGRATION INHIBITORY FACTOR (HUMAN); mRNA sequence.	Homo sapiens	32,489	10-Jun-99
		GB_IN2:AF153269	1308	AF153269	Tetrahymena thermophila dynein heavy chain (DYH10) gene, partial cds.	Tetrahymena thermophila	37,615	24-Jun-99
rx01470	549	GB_PL2:ATF3L17	94319	AL080283	Arabidopsis thaliana DNA chromosome 4, BAC clone F3L17 (ESSA project).	Arabidopsis thaliana	35,946	24-Jun-99
		GB_PL2:ATF3L17	94319	AL080283	Arabidopsis thaliana DNA chromosome 4, BAC clone F3L17 (ESSA project).	Arabidopsis thaliana	40,570	6-Aug-97
rx01471	930	GB_EST11:AA233898	460	AA233898	z149b12.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:866719 3', mRNA sequence.	Homo sapiens	36,277	6-Aug-97
		GB_EST11:AA234033	428	AA234033	z149b12.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:866719 5' similar to contains Alu repetitive element; mRNA sequence.	Homo sapiens	36,821	7-Jul-99
rx01472		GB_PL2:AF084971	1874	AF084971	Catharanthus roseus G-box binding protein 1 (GBF1) mRNA, complete cds.	Catharanthus roseus		
rx01473	888	GB_PL2:AC002329	76170	AC002329	DNA sequence of Arabidopsis thaliana BAC F5J6 from chromosome IV, complete sequence.	Arabidopsis thaliana	35,780	18-Nov-97
		GB_HTG3:AC009485	190706	AC009485	Homo sapiens clone NH0324G03, *** SEQUENCING IN PROGRESS ***; 6 unordered pieces.	Homo sapiens	39,649	04-OCT-1999
		GB_HTG3:AC009485	190706	AC009485	Homo sapiens clone NH0324G03, *** SEQUENCING IN PROGRESS ***; 6 unordered pieces.	Homo sapiens	39,649	04-OCT-1999
rx01474	669	GB_PR2:HSP373C6	85654	AL022393	Homo sapiens DNA sequence from P1 p373c6 on chromosome 6p21.31-21.33. Contains zinc finger proteins, pseudogenes, ESTs and STS.	Homo sapiens	33,934	22-Nov-99
		GB_EST4:H30893	485	H30893	yp43e11.r1 Soares retina N2b5HR Homo sapiens cDNA clone IMAGE:190220 5' similar to contains Alu repetitive element; mRNA sequence.	Homo sapiens	37,866	16-Aug-95
rx01475	549	GB_PL2:ATAC005398	80238	AC005398	Arabidopsis thaliana chromosome II BAC T6B13 genomic sequence, complete sequence.	Arabidopsis thaliana	38,640	30-OCT-1998
		GB_PL1:AB024028	70952	AB024028	Arabidopsis thaliana genomic DNA, chromosome 3, TAC clone: K1G2, complete sequence.	Arabidopsis thaliana	40,520	20-Nov-99
		GB_HTG2:AC007940	159279	AC007940	Homo sapiens clone 44_C_14, *** SEQUENCING IN PROGRESS ***; 12 unordered pieces.	Homo sapiens	36,044	1-Jul-99
		GB_HTG2:AC007940	159279	AC007940	Homo sapiens clone 44_C_14, *** SEQUENCING IN PROGRESS ***; 12 unordered pieces.	Homo sapiens	36,044	1-Jul-99
rx01476	465	GB_PL1:AB009087	1074	AB009087	Chlamydomonas sp. mRNA for alternative oxidase, partial cds.	Chlamydomonas sp.	38,596	05-DEC-1997
		GB_PL1:AB009087	1074	AB009087	Chlamydomonas sp. mRNA for alternative oxidase, partial cds.	Chlamydomonas sp.	39,519	05-DEC-1997
rx01479	363	GB_HTG1:HSJ136J15	148579	AL118496	Homo sapiens chromosome X clone RP1-136J15, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	39,886	23-Nov-99

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

GB_HTG1:HSJ136J15	148579	AL118496	Homo sapiens chromosome X clone RP1-136J15, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	39,886	23-Nov-99
GB_HTG2:AC007579	134344	AC007579	Drosophila melanogaster chromosome 2 clone BACR07M03 (D607) RP1-98 07 M.3 map 53A-53B strain Y, cn bw sp, *** SEQUENCING IN PROGRESS ***; 108 unordered pieces.	Drosophila melanogaster	39,266	2-Aug-99
GB_PR3:AF064858	193387	AF064858	Homo sapiens chromosome 21q22.3 BAC 28F9, complete sequence.	Homo sapiens	38,281	2-Jun-98
GB_GSS10:AQ207755	496	AQ207755	HS_3026_B1_G04_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3026 Col=7 Row=N, genomic survey sequence.	Homo sapiens	37,150	18-Sep-98
GB_PR3:AF064858	193387	AF064858	Homo sapiens chromosome 21q22.3 BAC 28F9, complete sequence.	Homo sapiens	38,379	2-Jun-98
GB_IN1:CEY50E8A	61864	AL117200	Caenorhabditis elegans cosmid Y50E8A, complete sequence.	Caenorhabditis elegans	39,617	19-Nov-99
GB_IN1:CEY50E8A	61864	AL117200	Caenorhabditis elegans cosmid Y50E8A, complete sequence.	Caenorhabditis elegans	34,349	19-Nov-99
GB_PL1:AB028606	61510	AB028606	Arabidopsis thaliana genomic DNA, chromosome 5, BAC clone:F16F17, complete sequence.	Arabidopsis thaliana	34,858	20-Nov-99
GB_BA1:CGPROMF10	60	X90358	C glutamicum DNA for promoter fragment F10.	Corynebacterium glutamicum	65,000	4-Nov-96
GB_EST10:AA183656	465	AA183656	mi20f08.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:621639 5' similar to WP:T02C12.2 CE01062.1; mRNA sequence.	Mus musculus	35,333	17-Feb-97
GB_EST9:AA110912	309	AA110912	mm02c01.r1 Stratagene mouse kidney (#937315) Mus musculus cDNA clone IMAGE:520320 5' similar to WP:T02C12.2 CE01062.1; mRNA sequence.	Mus musculus	37,662	4-Feb-97
GB_EST16:AA589576	567	AA589576	v49b06.s1 Stratagene mouse skin (#937313) Mus musculus cDNA clone IMAGE:975539 3' similar to gb:M63488 REPLICATION PROTEIN A 70 KD DNA-BINDING SUBUNIT (HUMAN); mRNA sequence.	Mus musculus	34,104	16-Sep-97
GB_EST28:AU051490	865	AU051490	AU051490 Sugano mouse brain mncb Mus musculus cDNA clone MNCb-2105 5', mRNA sequence.	Mus musculus	29,837	18-MAR-1999
GB_EST32:AI738234	609	AI738234	606044C01.x2 606 - Ear tissue cDNA library from Schmidt lab Zea mays cDNA, mRNA sequence.	Zea mays	38,291	16-Jun-99
GB_PR4:AC007106	172188	AC007106	Homo sapiens chromosome 4 clone C0383J20 map 4p16, complete sequence.	Homo sapiens	35,108	2-Jun-99
GB_BA2:AE001093	14097	AE001093	Archaeoglobus fulgidus section 14 of 172 of the complete genome.	Archaeoglobus fulgidus	38,113	15-DEC-1997
GB_PR4:AC007106	172188	AC007106	Homo sapiens chromosome 4 clone C0383J20 map 4p16, complete sequence.	Homo sapiens	34,657	2-Jun-99
GB_BA2:SCJ1	36925	AL109962	Streptomyces coelicolor cosmid J1.	Streptomyces coelicolor A3(2)	50,722	24-Sep-99
GB_BA1:MTY20B11	36330	Z95121	Mycobacterium tuberculosis H37Rv complete genome; segment 139/162.	Mycobacterium tuberculosis	40,238	17-Jun-98
GB_BA1:PAU12891	4062	U12891	Pseudomonas aeruginosa PAO substrain OT684 pyoverdine gene transcriptional regulator PvdS (pvdS) gene, complete cds.	Pseudomonas aeruginosa	39,856	03-OCT-1996
GB_PAT:178757	2203	I78757	Sequence 13 from patent US 5693781.	Unknown.	61,364	3-Apr-98
GB_PAT:192046	2203	I92046	Sequence 13 from patent US 5726299.	Unknown.	61,364	01-DEC-1998
GB_PR4:AC002427	101098	AC002427	Homo sapiens BAC clone GS011E15 from 5q31, complete sequence.	Homo sapiens	42,250	02-DEC-1998
GB_HTG3:AC008954	23046	AC008954	Homo sapiens chromosome 5 clone CITB-H1_2340N2, *** SEQUENCING IN PROGRESS ***; 51 unordered pieces.	Homo sapiens	37,830	3-Aug-99
GB_HTG3:AC008954	23046	AC008954	Homo sapiens chromosome 5 clone CITB-H1_2340N2, *** SEQUENCING IN PROGRESS ***; 51 unordered pieces.	Homo sapiens	37,830	3-Aug-99
GB_BA1:QPHQSOP	4068	L28041	Plasmid QpH1 (from Coxiella burnetii) qsoA and qsoB genes, promoter region.	Plasmid QpH1	37,517	14-Jul-95
GB_SY:SYNNM13GAL	66	M10216	Bacteriophage M13gt102 N-terminal beta-galactosidase gene (lac+ phenotype).	unidentified cloning vector	62,121	27-Apr-93
GB_SY:SYNNM13GAL	66	M10216	Bacteriophage M13gt102 N-terminal beta-galactosidase gene (lac+ phenotype).	unidentified cloning vector	62,121	27-Apr-93
GB_BA2:BUJ58817	2892	U58817	Bradyrhizobium japonicum aconitase (acnA) gene, complete cds.	Bradyrhizobium japonicum	40,214	13-Nov-96
GB_BA1:MTCY98	31225	Z83860	Mycobacterium tuberculosis H37Rv complete genome; segment 103/162.	Mycobacterium tuberculosis	38,274	17-Jun-98
GB_BA1:MTCY98	31225	Z83860	Mycobacterium tuberculosis H37Rv complete genome; segment 103/162.	Mycobacterium tuberculosis	33,710	17-Jun-98
GB_RO:D89731	1815	D89731	Rattus norvegicus mRNA for AIM-1, complete cds.	Rattus norvegicus	38,125	7-Feb-99

TABLE 4: ALIGNMENT RESULTS

GB_RO:D89731	1815	D89731	Rattus norvegicus mRNA for AIM-1, complete cds.	Rattus norvegicus	34,600	7-Feb-99
rx01518						
rx01519	870	GB_HTG2:AC007929	123885	AC007929	Drosophila melanogaster chromosome 3 clone BACR05C11 (D759) RPCI-98 05.C.11 map 95A-95C Drosophila melanogaster strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 85 unordered pieces.	2-Aug-99
		GB_HTG2:AC007929	123885	AC007929	Drosophila melanogaster chromosome 3 clone BACR05C11 (D759) RPCI-98 05.C.11 map 95A-95C Drosophila melanogaster strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 85 unordered pieces.	2-Aug-99
rx01520		GB_PL2:AC011437	95310	AC011437	Arabidopsis thaliana chromosome III BAC F7018 genomic sequence, complete sequence.	2-Nov-99
rx01523	1074	GB_GSS11:AQ270206	389	AQ270206	HS_2037_A1_G10_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2037 Col=19 Row=M, genomic survey sequence.	3-Nov-98
		GB_GSS14:AQ588624	480	AQ588624	CITBI-E1-2643C9.TF CITBI-E1 Homo sapiens genomic clone 2643C9, genomic survey sequence.	7-Jun-99
		GB_HTG1:CEY80D3	245017	AL020988	Caenorhabditis elegans chromosome V clone Y80D3, *** SEQUENCING IN PROGRESS *** in unordered pieces.	6-Sep-99
rx01525	1845	GB_PR3:HS560B9	99074	Z98751	Human DNA sequence from PAC 560B9 on chromosome 1q24-1q25. Contains profilin-like pseudogene, 60S ribosomal protein L4 pseudogene RNA binding protein, ESTs, GSS.	23-Nov-99
		GB_PR3:HS560B9	99074	Z98751	Human DNA sequence from PAC 560B9 on chromosome 1q24-1q25. Contains profilin-like pseudogene, 60S ribosomal protein L4 pseudogene RNA binding protein, ESTs, GSS.	23-Nov-99
rx01527						
rx01536	1365	GB_HTG3:AC010745	193862	AC010745	Homo sapiens clone NH0549D18, *** SEQUENCING IN PROGRESS ***. 30 unordered pieces.	21-Sep-99
		GB_HTG3:AC010745	193862	AC010745	Homo sapiens clone NH0549D18, *** SEQUENCING IN PROGRESS ***. 30 unordered pieces.	21-Sep-99
		GB_HTG3:AC010745	193862	AC010745	Homo sapiens clone NH0549D18, *** SEQUENCING IN PROGRESS ***. 30 unordered pieces.	21-Sep-99
rx01539	2289	GB_BA1:CGPROMIF10	60	X90358	C.glutamicum DNA for promoter fragment F10.	4-Nov-96
		GB_HTG3:AC011577	151996	AC011577	Homo sapiens clone 12_P_19, LOW-PASS SEQUENCE SAMPLING.	07-OCT-1999
		GB_HTG3:AC011577	151996	AC011577	Homo sapiens clone 12_P_19, LOW-PASS SEQUENCE SAMPLING.	07-OCT-1999
rx01540	825	GB_EST10:AA152819	491	AA152819	mq67e04.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:583806 5', mRNA sequence.	18-Feb-97
		GB_EST29:AI616027	481	AI616027	mq67e04.y1 Soares 2NbMT Mus musculus cDNA clone IMAGE:583806 5', mRNA sequence.	21-Apr-99
		GB_GSS6:AQ843942	488	AQ843942	LMAJFV1_Jm03d08.x2 Leishmania major FV1 random genomic library Leishmania major genomic clone LMAJFV1_Jm03d08 3', genomic survey sequence.	04-OCT-1999
rx01543	2889	GB_BA2:ECOUW67_0	110000	U18997	Escherichia coli K-12 chromosomal region from 67.4 to 76.0 minutes.	15-Sep-99
		GB_BA2:ECOUW67_0	110000	U18997	Escherichia coli K-12 chromosomal region from 67.4 to 76.0 minutes.	15-Sep-99
		GB_BA2:AE000397	14820	AE000397	Escherichia coli K-12 MG1655 section 287 of 400 of the complete genome.	12-Nov-98
rx01544	723	GB_EST37:AI977486	479	AI977486	EST272080 Schistosoma mansoni male, Phil LoVerde/Joel Merrick Schistosoma mansoni cDNA clone SMMA80 5' end, mRNA sequence.	27-Aug-99
		GB_VI:AF056119	564	AF056119	HIV-1 isolate z560 from Zimbabwe, envelope glycoprotein V3-V5 region (env) gene, partial cds.	26-Sep-98
		GB_VI:AF056119	564	AF056119	HIV-1 isolate z560 from Zimbabwe, envelope glycoprotein V3-V5 region (env) gene, partial cds.	26-Sep-98
rx01545	1374	GB_HTG1:CEY53C10_1110000	Z93340	Z93340	Caenorhabditis elegans chromosome I clone Y53C10, *** SEQUENCING IN PROGRESS *** in unordered pieces.	4-Nov-98
		GB_HTG1:CEY53C10_1110000	Z93340	Z93340	Caenorhabditis elegans chromosome I clone Y53C10, *** SEQUENCING IN PROGRESS *** in unordered pieces.	4-Nov-98

rxao1546	819	GB_HTG1:CEY47H9_1	110000	Z92853	GB_EST19:AA769027	473	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Caenorhabditis elegans	4-Sep-98	30,285
		GB_GSS3:B58207	420	B58207	GB_EST26:AU002373	837	AV025617	AV025617 Mus musculus adult C57BL/6J lung Mus musculus cDNA clone 1200014J18, mRNA sequence.	Homo sapiens	20-Jun-98	45,228
		GB_IN1:HAU02678	623	U02678	GB_EST19:AA769027	473	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Mitochondrion Helicoverpa armigera	27-OCT-1993	37,500
		GB_EST26:AU002373	837	AV025617	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Homo sapiens	28-Jan-98	39,583
		GB_EST30:AV025617	494	AV025617	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Bombyx mori	15-Jan-99	37,698
		GB_EST26:AU002373	837	AV025617	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Mus musculus	31-Aug-99	38,316
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Bombyx mori	15-Jan-99	36,767
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Plasmodium vivax	26-Apr-93	37,077
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Unknown.	07-OCT-1997	38,250
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Gallus gallus	15-OCT-1997	37,104
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Homo sapiens	11-Apr-99	39,401
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Panaxenus monodon	28-Jun-99	40,325
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Homo sapiens	27-Jan-95	38,868
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Homo sapiens	19-OCT-1999	36,296
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Homo sapiens	19-OCT-1999	36,296
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Homo sapiens	03-OCT-1999	36,440
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Rattus norvegicus	4-Jul-99	39,039
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Schizosaccharomyces pombe	17-Nov-99	36,991
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Hordeum vulgare	13-Jul-95	36,390
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Populus balsamifera subsp. trichocarpa	03-DEC-1998	40,326
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Arabidopsis thaliana	17-Aug-99	34,442
		GB_EST33:AI770254	555	AI770254	GB_EST26:AU002373	837	AA769027	oa78g01.s1 NCL_CGAB1 Homo sapiens cDNA clone IMAGE:1318416 3' similar to TR:Q92617 Q92617 MYELOBLAST KIA0220 .; mRNA sequence.	Arabidopsis thaliana	17-Aug-99	34,



**TABLE 4: ALIGNMENT RESULTS**

GB_GSS1:GGA300114	662	AJ231964	Gallus gallus anonymous sequence from Cosmid mapping to chicken chromosome 3 (Cosmid 30 - Contig 14), genomic survey sequence.	Gallus gallus	38,608	25-Aug-98
GB_PR4:AC004815	120538	AC004815	Homo sapiens clone 82F9, complete sequence.	Homo sapiens	35,383	8-Sep-99
GB_PR2:AP000151	100000	AP000151	Homo sapiens genomic DNA, chromosome 21q22.2, DSCR region, clone D47-S479, segment 3/16, complete sequence.	Homo sapiens	39,722	20-Nov-99
GB_PR2:AP000151	100000	AP000151	Homo sapiens genomic DNA, chromosome 21q22.2, DSCR region, clone D47-S479, segment 3/16, complete sequence.	Homo sapiens	34,683	20-Nov-99
GB_GSS5:AQ791385	522	AQ791385	HS_5268_A2_C05_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=844 Col=10 Row=E, genomic survey sequence.	Homo sapiens	43,235	3-Aug-99
GB_EST15:AA463774	255	AA463774	aa09f10.r1 Soares_NhhMPu_S1 Homo sapiens cDNA clone IMAGE:812779 5' similar to gb:X07868_ma1 PUTATIVE INSULIN-LIKE GROWTH FACTOR II ASSOCIATED (HUMAN); contains element PTRY repetitive element, mRNA sequence.	Homo sapiens	42,045	10-Jun-97
GB_PR1:AP000010	100000	AP000010	Homo sapiens genomic DNA of 21q22.2 Down Syndrome region, segment 3/13.	Homo sapiens	38,684	24-Apr-99
GB_HTG2:AC008163	135300	AC008163	Homo sapiens clone NH0005B09, *** SEQUENCING IN PROGRESS ***; 1 unordered pieces.	Homo sapiens	34,141	31-Jul-99
GB_HTG2:AC008163	135300	AC008163	Homo sapiens clone NH0005B09, *** SEQUENCING IN PROGRESS ***; 1 unordered pieces.	Homo sapiens	34,141	31-Jul-99
GB_HTG3:AC009231	181006	AC009231	Homo sapiens clone NH0350124, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	36,767	6-Aug-99
GB_BA1:MTIC1125	37432	Z98268	Mycobacterium tuberculosis H37Rv complete genome, segment 76/162.	Mycobacterium tuberculosis	57,955	17-Jun-98
GB_BA1:MTHVPROT	2544	X98295	M.tuberculosis TlyA gene.	Mycobacterium tuberculosis	57,955	2-Jun-98
GB_BA1:U00021	39193	U00021	Mycobacterium leprae cosmid L247.	Mycobacterium leprae	35,261	29-Sep-94
GB_HTG3:AC009462	103046	AC009462	Drosophila melanogaster chromosome 3 clone BACR27G04 (D985) RPCI-98 27 G.4 map 90E-90E strain y; cn bw sp, *** SEQUENCING IN PROGRESS***, 88 unordered pieces.	Drosophila melanogaster	38,898	30-Aug-99
GB_HTG3:AC009462	103046	AC009462	Drosophila melanogaster chromosome 3 clone BACR27G04 (D985) RPCI-98 27 G.4 map 90E-90E strain y; cn bw sp, *** SEQUENCING IN PROGRESS***, 88 unordered pieces.	Drosophila melanogaster	38,898	30-Aug-99
GB_HTG2:AC006889	267118	AC006889	Caenorhabditis elegans clone Y65B4, *** SEQUENCING IN PROGRESS ***; 6 unordered pieces.	Caenorhabditis elegans	36,236	26-Feb-99
GB_EST2:R02663	397	R02663	ye80a04.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:124014 5', mRNA sequence.	Homo sapiens	36,272	31-MAR-1995
GB_EST6:W03663	346	W03663	za65b08.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:297399 5', mRNA sequence.	Homo sapiens	42,775	19-Apr-96
GB_IN2:AC007765	163403	AC007765	Drosophila melanogaster, chromosome 2L, region 23C1-23C5, P1 clones DS02190 and DS00906, complete sequence.	Drosophila melanogaster	38,351	9-Jun-99
GB_STS:G39530	822	G39530	Z22942 Zebrafish AB Danio rerio STS genomic, sequence tagged site.	Danio rerio	41,242	31-Jul-98
GB_STS:G39530	822	G39530	Z22942 Zebrafish AB Danio rerio STS genomic, sequence tagged site.	Danio rerio	41,242	31-Jul-98
GB_PR3:HSDJ991C6	128995	AL078599	Human DNA sequence from clone 991C6 on chromosome 6q14.1-15, complete sequence.	Homo sapiens	35,188	23-Nov-99
GB_EST38:AW033000	685	AW033000	EST1276559 tomato callus, TAMU Lycopersicon esculentum cDNA clone cLEC19N6, mRNA sequence.	Lycopersicon esculentum	44,706	15-Sep-99
GB_PL2:SPAC3C7	35052	Z99568	S.pombe chromosome I cosmid c3C7.	Schizosaccharomyces pombe	38,158	29-Nov-99
GB_PL2:SPAC3C7	35052	Z99568	S.pombe chromosome I cosmid c3C7.	Schizosaccharomyces pombe	37,888	29-Nov-99
GB_HTG3:AC008689	145122	AC008689	Homo sapiens chromosome 5 clone CIT978SKB_61G23, *** SEQUENCING IN PROGRESS ***; 44 unordered pieces.	Homo sapiens	35,448	3-Aug-99
GB_HTG3:AC008689	145122	AC008689	Homo sapiens chromosome 5 clone CIT978SKB_61G23, *** SEQUENCING IN PROGRESS ***; 44 unordered pieces.	Homo sapiens	35,448	3-Aug-99

TABLE 4: ALIGNMENT RESULTS

rx01610	852	GB_HTG3:AC009175	233932	AC009175	Homo sapiens chromosome 5 clone RPCI-PAC_241C15, *** SEQUENCING IN PROGRESS ***	94	Homo sapiens	37,423	3-Aug-99
		GB_BA1:ECU82664	139818	U82664	unordered pieces.				
		GB_BA1:ECU82664	139818	U82664	Escherichia coli minutes 9 to 11 genomic sequence.		Escherichia coli	34,198	11-Jan-97
		GB_BA1:D90845	139818	U82664	Escherichia coli minutes 9 to 11 genomic sequence.		Escherichia coli	36,949	11-Jan-97
		GB_PR3:AC005581	19366	D90845	E.coli genomic DNA, Kohara clone #356(46.1-46.5 min.).		Escherichia coli	39,808	21-MAR-1997
rx01611	480	GB_EST35:AI809560	39601	AC005581	Homo sapiens chromosome 19, cosmid R31237, complete sequence.		Homo sapiens	38,901	1-Sep-98
		GB_EST35:AI809560	443	AI809560	wf30h09.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2357153 3', mRNA sequence.		Homo sapiens	35,989	7-Jul-99
		GB_EST25:AI279953	416	AI279953	qh9b02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1854219 3', mRNA sequence.		Homo sapiens	35,276	27-Jan-99
rx01618	507	GB_GSS14:AQ506815	616	AQ506815	RPCI-11-316B19, TV RPCI-11 Homo sapiens genomic clone RPCI-11-316B19, genomic survey sequence.		Homo sapiens	41,901	29-Apr-99
		GB_HTG2:AC007512	168140	AC007512	Homo sapiens clone hRPK.96_A_1, *** SEQUENCING IN PROGRESS ***	14	unordered pieces.	39,549	07-MAY-1999
		GB_HTG2:AC007512	168140	AC007512	Homo sapiens clone hRPK.96_A_1, *** SEQUENCING IN PROGRESS ***	14	unordered pieces.	39,549	07-MAY-1999
rx01619	1098	GB_BA1:ECOUW93	338534	U14003	Escherichia coli K-12 chromosomal region from 92.8 to 00.1 minutes.		Escherichia coli	39,459	17-Apr-96
		GB_BA2:AE000497	10334	AE000497	Escherichia coli K-12 MG-1655 section 387 of 400 of the complete genome.		Escherichia coli	39,459	12-Nov-98
		GB_BA1:ECOLEUX	2147	M96355	Escherichia coli leucine tRNA gene and ORF1 gene, complete cds.		Escherichia coli	41,434	26-Jul-95
rx01622	726	GB_PR4:AC007664	162470	AC007664	Homo sapiens, complete sequence.		Homo sapiens	37,006	18-Sep-99
		GB_PR4:AC002472	147100	AC002472	Homo sapiens Chromosome 22q11.2 PAC Clone p_n5 In BCRL2-GGT Region, complete sequence.		Homo sapiens	33,058	13-Sep-99
rx01623	834	GB_BA1:SYCSLRG	135638	D64005	Synechocystis sp. PCC6803 complete genome. 24/27. 3002966-3138603		Synechocystis sp.	38,755	13-Feb-99
		GB_HTG4:AC007422	156484	AC007422	Homo sapiens clone hRPK.68_A_1, *** SEQUENCING IN PROGRESS ***	3	unordered pieces.	36,759	23-OCT-1999
		GB_HTG4:AC007422	156484	AC007422	Homo sapiens clone hRPK.68_A_1, *** SEQUENCING IN PROGRESS ***	3	unordered pieces.	36,759	23-OCT-1999
		GB_HTG4:AC007422	156484	AC007422	Homo sapiens clone hRPK.68_A_1, *** SEQUENCING IN PROGRESS ***	3	unordered pieces.	33,495	23-OCT-1999
rx01624	468	GB_BA2:AE001816	10007	AE001816	Thermotoga maritima section 128 of 136 of the complete genome.		Thermotoga maritima	38,852	2-Jun-99
		GB_RO:CCPH20	2152	X56332	Cavia cobaya mRNA for PH-20 protein.		Cavia porcellus	38,293	20-Feb-91
		GB_PAT:189388	2152	I89388	Sequence 1 from patent US 5721348.		Unknown.	38,293	10-Aug-98
rx01628	1137	GB_BA1:CGPROPGEN	2936	Y12537	C.glutamicum proP gene.		Corynebacterium glutamicum	37,882	17-Nov-98
		GB_IN1:CET06E6	33238	Z81117	Caenorhabditis elegans cosmid T06E6, complete sequence.		Caenorhabditis elegans	38,330	2-Sep-99
		GB_HTG2:AC006916	157093	AC006916	Caenorhabditis elegans clone Y9C9, *** SEQUENCING IN PROGRESS ***	3	unordered pieces.	36,519	24-Feb-99
rx01630	1200	GB_PR3:AC004945	167372	AC004945	Homo sapiens PAC clone DJ097N05 from 7q11.23-q21.1, complete sequence.		Homo sapiens	34,996	26-Sep-98
		GB_PR4:AC006385	173508	AC006385	Homo sapiens clone NH0559J05, complete sequence.		Homo sapiens	32,525	05-MAY-1999
		GB_PR3:AC004945	167372	AC004945	Homo sapiens PAC clone DJ097N05 from 7q11.23-q21.1, complete sequence.		Homo sapiens	37,972	26-Sep-98
rx01634	675	GB_PR4:AC005901	156763	AC005901	Homo sapiens chromosome 17, clone hRPK.15_K_2, complete sequence.		Homo sapiens	35,736	20-Jan-99
		GB_PR4:AC005901	156763	AC005901	Homo sapiens chromosome 17, clone hRPK.15_K_2, complete sequence.		Homo sapiens	35,549	20-Jan-99
rx01635	864	GB_HTG3:AC008590	268470	AC008590	Homo sapiens chromosome 5 clone CIT-HSPC_575D19, *** SEQUENCING IN PROGRESS ***		Homo sapiens	35,370	3-Aug-99
		GB_HTG3:AC008590	268470	AC008590	235 unordered pieces.		Homo sapiens	35,370	3-Aug-99
		GB_HTG3:AC008590	268470	AC008590	235 unordered pieces.		Homo sapiens	38,571	3-Aug-99
rx01638	1023	GB_HTG3:AC008261	93735	AC008261	Arabidopsis thaliana chromosome III clone TAMU-T4P13, *** SEQUENCING IN PROGRESS ***	1	Arabidopsis thaliana	37,200	10-Aug-99

TABLE 4: ALIGNMENT RESULTS

GB_HTG3:AC008261	93735	AC008261	Arabidopsis thaliana chromosome III clone TAMU-T4P13, *** SEQUENCING IN PROGRESS ***	1	Arabidopsis thaliana	37,200	10-Aug-99
GB_HTG3:AC008261	93735	AC008261	ordered pieces.				
GB_HTG3:AC008261	93735	AC008261	Arabidopsis thaliana chromosome III clone TAMU-T4P13, *** SEQUENCING IN PROGRESS ***	1	Arabidopsis thaliana	32,637	10-Aug-99
GB_PR3:HS881N20	130263	AL031670	ordered pieces.				
GB_PR3:HS881N20	130263	AL031670	Human DNA sequence from clone 681N20 on chromosome 20p12.1-13 Contains FTLL1(ferritin, light polypeptide-like 1), a gene similar to Zinc finger, C3HC4 type (RING finger), weakly similar to SW:GOLL_DROME_Q06003 GOLITH PROTEIN, ESTs, STS, GSS, CA repeat (D20S889), CpG Islands, complete sequence.		Homo sapiens	34,676	23-Nov-99
GB_HTG4:AC011177	168660	AC011177	Homo sapiens clone 11_L_8, *** SEQUENCING IN PROGRESS ***	7	unordered pieces.		
GB_HTG4:AC011177	168660	AC011177	Homo sapiens clone 11_L_8, *** SEQUENCING IN PROGRESS ***	7	unordered pieces.		
GB_BA1:MTY15F10	38204	Z94121	Mycobacterium tuberculosis H37Rv complete genome; segment 161/162.		Mycobacterium tuberculosis	37,723	29-OCT-1999
GB_BA1:SC3C3	31382	AL031231	Streptomyces coelicolor cosmid 3C3.		Streptomyces coelicolor	39,602	10-Aug-98
GB_BA1:SCU33176	1238	U33176	Streptomyces coelicolor serine protease gene, complete cds.		Streptomyces coelicolor	41,467	9-Jan-96
GB_IN1:CEZC101	26147	Z93395	Caenorhabditis elegans cosmid ZC101, complete sequence.		Caenorhabditis elegans	36,280	23-Jul-99
GB_EST16:C47058	360	C47058	C47058 Yuji Kohara unpublished cDNA: Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone Yk432e11 5', mRNA sequence.		Caenorhabditis elegans	39,944	18-OCT-1999
GB_HTG1:CEY54E2_1	110000	Z92861	Caenorhabditis elegans chromosome II clone Y54E2, *** SEQUENCING IN PROGRESS ***	in	Caenorhabditis elegans	36,280	29-Jul-98
GB_BA1:MTV025	121125	AL022121	unordered pieces.				
GB_BA1:MTV025	121125	AL022121	Mycobacterium tuberculosis H37Rv complete genome; segment 155/162.		Mycobacterium tuberculosis	60,310	24-Jun-99
GB_BA1:MLCB2407	35615	AL023596	Mycobacterium leprae cosmid B2407.		Mycobacterium leprae	37,378	27-Aug-99
GB_BA1:MSGB577COS37770	L01263	L01263	M. leprae genomic dna sequence, cosmid b577.		Mycobacterium leprae	57,829	14-Jun-96
GB_BA1:MSGB577COS37770	L01263	L01263	M. leprae genomic dna sequence, cosmid b577.		Mycobacterium leprae	53,781	14-Jun-96
GB_BA1:MLCB2407	35615	AL023596	Mycobacterium leprae cosmid B2407.		Mycobacterium leprae	39,348	27-Aug-99
GB_BA1:MTV025	121125	AL022121	Mycobacterium tuberculosis H37Rv complete genome; segment 155/162.		Mycobacterium tuberculosis	53,503	24-Jun-99
GB_GSS1:CNS010VB	837	AL099473	Drosophila melanogaster genome survey sequence T7 end of BAC BACN05B20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.		Drosophila melanogaster	39,412	26-Jul-99
GB_EST1:T55021	429	T55021	yb42e09.r1 Stralagene fetal spleen (#937205) Homo sapiens cDNA clone IMAGE:73864 5', mRNA sequence.		Homo sapiens	39,730	6-Feb-95
GB_EST38:AL120803	505	AL120803	DKFZp762E202_r1 762 (synonym: hmcl2) Homo sapiens cDNA clone DKFZp762E202 5', mRNA sequence.		Homo sapiens	40,476	27-Sep-99
GB_BA1:MTVCY1A10	25949	Z95387	Mycobacterium tuberculosis H37Rv complete genome; segment 117/162.		Mycobacterium tuberculosis	37,771	17-Jun-98
GB_BA1:MLCL581	36225	Z96801	Mycobacterium leprae cosmid L581.		Mycobacterium leprae	40,041	24-Jun-97
GB_PL2:ATAC003028	106448	AC003028	Arabidopsis thaliana chromosome II BAC F16M14 genomic sequence, complete sequence.		Arabidopsis thaliana	57,692	22-Jul-98
GB_BA1:MTV014	58280	AL021646	Mycobacterium tuberculosis H37Rv complete genome; segment 137/162.		Mycobacterium tuberculosis	38,877	18-Jun-98
GB_BA1:MTV030	29256	AL021428	Mycobacterium tuberculosis H37Rv complete genome; segment 4/162.		Mycobacterium tuberculosis	57,048	17-Jun-98
GB_BA1:MTV014	58280	AL021646	Mycobacterium tuberculosis H37Rv complete genome; segment 137/162.		Mycobacterium tuberculosis	40,511	18-Jun-98
GB_HTG3:AC008604	136016	AC008604	Homo sapiens chromosome 5 clone CIT978SKB_109F8, *** SEQUENCING IN PROGRESS ***	79	Homo sapiens	35,106	3-Aug-99
GB_HTG3:AC008604	136016	AC008604	unordered pieces.				
GB_HTG3:AC008604	136016	AC008604	Homo sapiens chromosome 5 clone CIT978SKB_109F8, *** SEQUENCING IN PROGRESS ***	79	Homo sapiens	35,106	3-Aug-99
GB_HTG3:AC011403	124903	AC011403	Homo sapiens chromosome 5 clone CIT978SKB_3P13, *** SEQUENCING IN PROGRESS ***	7	Homo sapiens	33,997	06-OCT-1999
GB_HTG2:AC008799	278007	AC008799	Caenorhabditis elegans clone Y51H7, *** SEQUENCING IN PROGRESS ***	7	unordered pieces.		
GB_HTG2:AC008799	278007	AC008799	Caenorhabditis elegans		Caenorhabditis elegans	34,787	23-Feb-99

TABLE 4: ALIGNMENT RESULTS

GB_HTG2:AC006799	278007	AC006799	Caenorhabditis elegans clone Y51H7, *** SEQUENCING IN PROGRESS ***	7 unordered pieces.	Caenorhabditis elegans	34,787	23-Feb-99
GB_HTG2:AC006799	278007	AC006799	Caenorhabditis elegans clone Y51H7, *** SEQUENCING IN PROGRESS ***	7 unordered pieces.	Caenorhabditis elegans	39,148	23-Feb-99
GB_BA1:CGU43535	2531	U43535	Corynebacterium glutamicum multidrug resistance protein (cmr) gene, complete cds.		Corynebacterium glutamicum	37,353	9-Apr-97
GB_IN2:AF062971	359	AF062971	Anoploleptus portus 28S ribosomal RNA gene, partial sequence.		Anoploleptus portus	42,462	20-Apr-99
GB_EST31:AU062109	544	AU062109	AU062109 Dictyostelium discoideum SL (H1 Urushihara) Dictyostelium discoideum cDNA clone SLH629, mRNA sequence.		Dictyostelium discoideum	34,810	20-MAY-1999
GB_BA2:CGU89648	1105	U89648	Corynebacterium glutamicum unidentified sequence involved in histidine biosynthesis, partial sequence.		Corynebacterium glutamicum	65,152	30-MAR-1999
GB_IN1:DMC86E4	29352	AL021086	Drosophila melanogaster cosmid clone 86E4.		Drosophila melanogaster	39,289	27-Apr-99
GB_PL1:ENY13759	3141	Y13759	Emicella nidulans abfB gene.		Emicella nidulans	38,552	05-MAR-1999
GB_PH:MPU46938	15664	U46938	Mycobacterium phage DS6A, Spe1/NheI G fragment sequence.		Mycobacterium phage DS6A	60,751	29-Jun-96
GB_PAT:J31047	15664	J31047	Sequence 3 from patent US 5582969.		Unknown.	60,751	6-Feb-97
GB_PAT:J36863	15664	J36863	Sequence 3 from patent US 5612182.		Unknown.	60,751	13-MAY-1997
GB_GSS14:AAQ512046	462	AQ512046	HS_5113_B1_B10_SPE6 RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=689 Col=19 Row=D, genomic survey sequence.		Homo sapiens	40,882	05-MAY-1999
GB_GSS13:AAQ475012	253	AQ475012	CITBI-E1-2592A3 TR CITBI-E1 Homo sapiens genomic clone 2592A3, genomic survey sequence.		Homo sapiens	35,417	23-Apr-99
GB_GSS14:AAQ512046	462	AQ512046	HS_5113_B1_B10_SPE6 RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=689 Col=19 Row=D, genomic survey sequence.		Homo sapiens	35,106	05-MAY-1999
GB_HTG5:AC010988	176452	AC010988	Homo sapiens clone NH0570F04, WORKING DRAFT SEQUENCE, 1 unordered pieces.		Homo sapiens	37,489	11-Nov-99
GB_HTG5:AC010988	176452	AC010988	Homo sapiens clone NH0570F04, WORKING DRAFT SEQUENCE, 1 unordered pieces.		Homo sapiens	35,492	11-Nov-99
GB_EST19:AA797620	497	AA797620	w26b03 r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone IMAGE:1244909		Mus musculus	44,318	10-Feb-98
GB_IN2:AF076597	846	AF076597	5', mRNA sequence.		Dictyostelium discoideum	37,149	30-Jul-98
GB_EST22:C94466	698	C94466	Dictyostelium discoideum ubiquitin-conjugating enzyme protein UbcC (ubcC) mRNA, complete cds.		Dictyostelium discoideum	33,647	15-Jun-98
GB_IN1:CELRO3H10	35080	U29382	C94466 Dictyostelium discoideum SS (H1 Urushihara) Dictyostelium discoideum cDNA clone SSL136, mRNA sequence.		Caenorhabditis elegans	40,951	22-Jun-95
GB_EST37:AI998397	584	AI998397	Caenorhabditis elegans cosmid R03H10.		Arabidopsis thaliana	37,333	8-Sep-99
GB_GSS14:AAQ551731	577	AQ551731	701545552 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis thaliana cDNA clone 701545552, mRNA sequence.		Homo sapiens	39,321	28-MAY-1999
GB_PL2:ATF618	122322	AL022198	RPCI-11-383J4 TJ RPCI-11 Homo sapiens genomic clone RPCI-11-383J4, genomic survey sequence.		Arabidopsis thaliana	34,335	27-Aug-99
GB_PL2:ATF618	122322	AL022198	Arabidopsis thaliana DNA chromosome 4, BAC clone F6118 (ESSA project).		Arabidopsis thaliana	34,895	27-Aug-99
GB_OV:XLSDK2	2306	Y10350	Arabidopsis thaliana DNA chromosome 4, BAC clone F6118 (ESSA project).		Xenopus laevis	36,538	2-Apr-98
GB_HTG3:AC010798	172575	AC010798	X laevis mRNA for nuclear protein SDK2.		Homo sapiens	37,116	23-Sep-99
GB_HTG3:AC010798	172575	AC010798	Homo sapiens chromosome 18 clone 470_B_24 map 18, *** SEQUENCING IN PROGRESS ***	7 unordered pieces.	Homo sapiens	37,116	23-Sep-99
GB_PL2:AF085231	3158	AF085231	Homo sapiens chromosome 18 clone 470_B_24 map 18, *** SEQUENCING IN PROGRESS ***	7 unordered pieces.	Arabidopsis thaliana	42,149	3-Jun-99
GB_HTG1:CEY48A6	296699	Z92854	Arabidopsis thaliana phytochelatin synthase 1 (APCS1) gene, complete cds.		Caenorhabditis elegans	37,393	31-Jul-98
GB_HTG1:CEY48A6	296699	Z92854	Caenorhabditis elegans chromosome III clone Y48A6, *** SEQUENCING IN PROGRESS ***	in unordered pieces.	Caenorhabditis elegans	37,393	31-Jul-98
GB_HTG1:CEY48A6	296699	Z92854	Caenorhabditis elegans chromosome III clone Y48A6, *** SEQUENCING IN PROGRESS ***	in unordered pieces.	Caenorhabditis elegans	32,946	31-Jul-98

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

rx01686	564	GB_EST35:AV156415	240	AV156415	AV156415 Mus musculus head C57BL/6J 12-day embryo Mus musculus cDNA clone 3000002G04, mRNA sequence.	Mus musculus	40,417	7-Jul-99
		GB_GSS15:AQ611518	597	AQ611518	HS_5084_B2_H01_T7A RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=650 Col=2 Row=P, genomic survey sequence.	Homo sapiens	41,848	15-Jun-99
		GB_BA1:PDIORAB	5467	Z48918	P diminuta iorA and iorB genes for isosquiline 1-oxidoreductase.	Brevundimonas diminuta	36,709	15-Aug-95
rx01693	1545	GB_IN1:CEF4686	36616	Z70780	Caenorhabditis elegans cosmid F4686, complete sequence.	Caenorhabditis elegans	36,856	04-DEC-1998
		GB_EST5:L44364	362	L44364	HUMEST1G5 Human thymus NSTH II Homo sapiens cDNA, mRNA sequence.	Homo sapiens	38,136	17-Jan-96
		GB_PR2:AF009282	414	AF009282	Homo sapiens clone FB3 Crl-du-chat region mRNA.	Homo sapiens	46,809	17-Aug-97
rx01694	1023	GB_BA1:CGA224946	2408	AJ224946	Corynebacterium glutamicum DNA for L-Malate:quinone oxidoreductase.	Corynebacterium glutamicum	100,000	11-Aug-98
		GB_EST7:W22650	715	W22650	71B2 Human retina cDNA Tsp509I-cleaved sublibrary Homo sapiens cDNA not directional, mRNA sequence.	Homo sapiens	37,538	06-MAY-1996
		GB_PR4:AC005343	137213	AC005343	Homo sapiens Chromosome 12p13.3 BAC RPC111-21K20 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	Homo sapiens	38,048	2-Apr-99
rx01696	1302	GB_BA1:MSGB1529CS	36985	L78824	Mycobacterium leprae cosmid B1529 DNA sequence.	Mycobacterium leprae	67,442	15-Jun-96
		GB_BA1:SC6A5	43632	AL049485	Streptomyces coelicolor cosmid 6A5.	Streptomyces coelicolor	66,821	24-MAR-1999
		GB_BA1:MTV003	13246	AL008883	Mycobacterium tuberculosis H37Rv complete genome; segment 125/162.	Mycobacterium tuberculosis	38,802	17-Jun-98
rx01697	1080	GB_PL2:HNNHAHR	2559	L76588	Helianthus annuus homeodomain protein 1 mRNA, complete cds.	Helianthus annuus	38,257	02-OCT-1997
		GB_PL2:HNNHAHR	2559	L76588	Helianthus annuus homeodomain protein 1 mRNA, complete cds.	Helianthus annuus	37,196	02-OCT-1997
rx01701	472	GB_BA2:AF069748	2103	AF069748	Pseudomonas chlororaphis polyurethanase esterase A (pueA) gene, complete cds.	Pseudomonas chlororaphis	39,130	4-Apr-99
		GB_BA2:AF069748	2103	AF069748	Pseudomonas chlororaphis polyurethanase esterase A (pueA) gene, complete cds.	Pseudomonas chlororaphis	40,271	4-Apr-99
rx01703	1236	GB_BA1:CGFDA	3371	X17313	Corynebacterium glutamicum fda gene for fructose-bisphosphate aldolase (EC 4.1.2.13).	Corynebacterium glutamicum	100,000	12-Sep-93
		GB_EST18:AA728419	340	AA728419	33598 CD4-6 Arabidopsis thaliana cDNA clone K2C9RP, mRNA sequence.	Arabidopsis thaliana	36,176	5-Jan-98
		GB_PL2:ATU90439	93639	U90439	Arabidopsis thaliana chromosome II BAC T06D20 genomic sequence, complete sequence.	Arabidopsis thaliana	35,381	21-Jul-97
rx01709	555	GB_OV:AF033670	2353	AF033670	Gallus gallus T-Box protein 4 (TBX4) mRNA, complete cds.	Gallus gallus	37,917	9-Jun-98
		GB_EST5:H85635	533	H85635	ys88c04.r1 Soares retina N2b5HR Homo sapiens cDNA clone IMAGE:221862 5', mRNA sequence.	Homo sapiens	36,735	14-Nov-95
rx01711	1281	GB_EST35:A1829867	427	A1829867	wj58c05.x1 NCLCGAP_Lu19 Homo sapiens cDNA clone IMAGE:2407016 3', mRNA sequence.	Homo sapiens	40,299	26-Aug-99
		GB_GSS10:AQ217798	441	AQ217798	HS_2007_A1_A04_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2007 Col=7 Row=A, genomic survey sequence.	Homo sapiens	38,292	19-Sep-98
		GB_PAT:E05055	2053	E05055	DNA encoding D-threonine aldolase(DTA).	Xanthomonas oryzae	36,475	29-Sep-97
		GB_PR3:HSJ570L12	143508	AL049589	Human DNA sequence from clone 570L12 on chromosome Xq13.1-21.1. Contains the PGK1 gene for phosphoglycerate kinase 1, the gene for a novel protein similar to TAF2G (TATA box binding protein (TBP)-associated factor, RNA polymerase II, G, 32kD) (TAFI31)). ESTs, STS, GSSs and a putative CpG island, complete sequence.	Homo sapiens	36,262	23-Nov-99
rx01714								
rx01715	819	GB_BA1:MTCY441	35187	Z80225	Mycobacterium tuberculosis H37Rv complete genome; segment 118/162.	Mycobacterium tuberculosis	36,216	18-Jun-98
		GB_BA1:SC1B5	14866	AL023517	Streptomyces coelicolor cosmid 1B5.	Streptomyces coelicolor	56,807	11-MAY-1998
		GB_BA1:MSGB1912CS	38503	L01536	M. leprae genomic dna sequence, cosmid b1912.	Mycobacterium leprae	55,651	14-Jun-96
rx01729	642	GB_BA2:CORCSLYS	2821	M89931	Corynebacterium glutamicum beta C-S lyase (aecD) and branched-chain amino acid uptake carrier (brnQ) genes, complete cds, and hypothetical protein YhbW (yhbW) gene, partial cds.	Corynebacterium glutamicum	41,993	4-Jun-98

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

GB_HTG4:AC010037	166249	AC010037	Drosophila melanogaster chromosome 3L/66B8 clone RPC198-6E4, *** SEQUENCING IN PROGRESS ***; 52 unordered pieces.	Drosophila melanogaster	34,951	16-OCT-1999
GB_HTG4:AC010037	166249	AC010037	Drosophila melanogaster chromosome 3L/66B8 clone RPC198-6E4, *** SEQUENCING IN PROGRESS ***; 52 unordered pieces.	Drosophila melanogaster	34,951	16-OCT-1999
GB_GSS3:B83621	616	B83621	RPC111-16H17.TPB RPC1-11 Homo sapiens genomic clone RPC1-11-16H17, genomic survey sequence.	Homo sapiens	40,899	9-Apr-99
GB_GSS3:B82563	586	B82563	RPC111-16N17.TP RPC1-11 Homo sapiens genomic clone RPC1-11-16N17, genomic survey sequence.	Homo sapiens	40,235	9-Apr-99
GB_HTG3:AC009074	45999	AC009074	Homo sapiens chromosome 16 clone RPC1-11_323C21, *** SEQUENCING IN PROGRESS ***; 27 unordered pieces.	Homo sapiens	36,783	3-Aug-99
GB_GSS14:AQ545618	327	AQ545618	CITB1-E1-2636O13. TR CITB1-E1 Homo sapiens genomic clone 2636O13, genomic survey sequence.	Homo sapiens	38,154	28-MAY-1999
GB_HTG2:AC006629	27659	AC006629	Caenorhabditis elegans clone F12E12, *** SEQUENCING IN PROGRESS ***; 1 unordered pieces.	Caenorhabditis elegans	35,658	23-Feb-99
GB_HTG2:AC006629	27659	AC006629	Caenorhabditis elegans clone F12E12, *** SEQUENCING IN PROGRESS ***; 1 unordered pieces.	Caenorhabditis elegans	35,714	23-Feb-99
GB_HTG1:CNS01DSB	222193	AL121768	Homo sapiens chromosome 14 clone R-976B16, *** SEQUENCING IN PROGRESS ***; in ordered pieces.	Homo sapiens	36,978	05-OCT-1999
GB_HTG1:CNS01DSB	222193	AL121768	Homo sapiens chromosome 14 clone R-976B16, *** SEQUENCING IN PROGRESS ***; in ordered pieces.	Homo sapiens	36,978	05-OCT-1999
GB_HTG5:AC011170	171788	AC011170	Homo sapiens clone 10_L_13, *** SEQUENCING IN PROGRESS ***; 14 unordered pieces.	Homo sapiens	35,259	5-Nov-99
GB_JN1:CEK04G11	34190	Z78544	Caenorhabditis elegans cosmid K04G11, complete sequence.	Caenorhabditis elegans	35,055	6-Sep-99
GB_HTG2:AC005995	170023	AC005995	Homo sapiens clone DJ0042M02, *** SEQUENCING IN PROGRESS ***; 13 unordered pieces.	Homo sapiens	33,624	23-Apr-99
GB_HTG2:AC005995	170023	AC005995	Homo sapiens clone DJ0042M02, *** SEQUENCING IN PROGRESS ***; 13 unordered pieces.	Homo sapiens	33,624	23-Apr-99
GB_EST30:AI658116	497	AI658116	fc22e07.y1 Zebrafish WashU MPMING EST Danio rerio cDNA 5' similar to TR.Q15883 Q15883 X104.1, mRNA sequence.	Danio rerio	35,604	06-MAY-1999
GB_HTG1:CEY60A9	275370	AL022281	Caenorhabditis elegans chromosome X clone Y60A9, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	35,610	2-Sep-99
GB_HTG1:CEY60A9	275370	AL022281	Caenorhabditis elegans chromosome X clone Y60A9, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	35,610	2-Sep-99
GB_GSS13:AQ447364	439	AQ447364	mgxb0006N01r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0006N01r, genomic survey sequence.	Magnaporthe grisea	34,889	8-Apr-99
GB_PR2:HSJ247G22	120487	AL096866	Human DNA sequence from clone RP1-247G22 on chromosome 6p12.2-21.2, complete sequence.	Homo sapiens	39,310	22-Nov-99
GB_GSS5:AQ811603	375	AQ811603	HS_5460_B1_C09_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=1036 Col=17 Row=F, genomic survey sequence.	Homo sapiens	38,133	25-Aug-99
GB_HTG5:AC009544	211057	AC009544	Homo sapiens chromosome 11 clone RP11-85D24 map 11, WORKING DRAFT SEQUENCE, 14 unordered pieces.	Homo sapiens	37,433	19-Nov-99
GB_HTG5:AC009544	211057	AC009544	Homo sapiens chromosome 11 clone RP11-85D24 map 11, WORKING DRAFT SEQUENCE, 14 unordered pieces.	Homo sapiens	37,069	19-Nov-99
GB_PR1:HUMPROFILX	17630	M96943	Human profilaggrin gene exons 1-3, 5' end.	Homo sapiens	36,352	27-Apr-93
GB_PL2:AF049174	535	AF049174	Tolypocladium inflatum NRRL 28024 28S ribosomal RNA gene, partial sequence.	Tolypocladium inflatum	40,187	03-MAR-1999
GB_BA1:MLCL458	43839	AL049478	Mycobacterium leprae cosmid L458.	Mycobacterium leprae	40,322	27-Aug-99
GB_PR1:HUMPROFILX	17630	M96943	Human profilaggrin gene exons 1-3, 5' end.	Homo sapiens	36,794	27-Apr-93
GB_PL1:YSCFBPA	2202	JO3207	S.cerevisiae fructose-1,6-bisphosphatase (FBP) gene, complete cds.	Saccharomyces cerevisiae	37,762	27-Apr-93

# TABLE 4: ALIGNMENT RESULTS

mosome XII cosmid 8039.

Saccharomyces cerevisiae	37,685	22-Aug-97
Caenorhabditis elegans	40,945	16-Nov-95
Caenorhabditis elegans	35,484	16-Nov-95
Mus musculus	35,264	4-Feb-99
Homo sapiens	37,037	11-MAR-1999
Homo sapiens	37,037	3-Jun-98
Homo sapiens	35,995	11-MAR-1999
Homo sapiens	39,566	3-Jun-99
Mus musculus	37,353	5-Sep-96
Caenorhabditis elegans	39,590	2-Nov-95
Caenorhabditis elegans	36,667	2-Sep-99
Caenorhabditis elegans	35,094	2-Sep-99
Caenorhabditis elegans	33,276	7-Sep-99
Drosophila melanogaster	35,740	16-OCT-1999
Drosophila melanogaster	35,740	16-OCT-1999
Caenorhabditis elegans	37,870	8-Aug-94
Strongylocentrotus purpuratus	39,337	21-Jul-94
Homo sapiens	39,626	24-MAY-1999
Homo sapiens	38,279	10-Apr-98
Corynebacterium glutamicum	100,000	17-Nov-98
Lycopersicon esculentum	41,187	29-Jun-99
Caenorhabditis elegans	38,077	23-Nov-98
Homo sapiens	42,142	09-DEC-1998
Homo sapiens	42,142	09-DEC-1998
Homo sapiens	37,748	09-DEC-1998
Homo sapiens	38,054	23-Nov-99

TABLE 4: ALIGNMENT RESULTS

rx01775	687	GB_HTG1: HSA557H15	228174	AL078590	Homo sapiens chromosome 6 clone RP11-557H15, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Homo sapiens	38,054	23-Nov-99
		GB_PL2: YSCH9998	35600	U00030	Saccharomyces cerevisiae chromosome VIII cosmid 9998.	Saccharomyces cerevisiae	39,110	2-Sep-97
		GB_PR3: HS37J18	131427	Z98747	Human DNA sequence from clone 37J18 on chromosome 1p36.2-36.3. Contains a putative novel gene, ESTs and GSSs, complete sequence.	Homo sapiens	35,036	23-Nov-99
		GB_HTG2: AC007475	185087	AC007475	Drosophila melanogaster chromosome 2 clone BACR04E21 (D592) RPCI-98 04.E.21 map 49A-49B strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***, 36 unordered pieces.	Drosophila melanogaster	37,333	2-Aug-99
		GB_HTG2: AC007475	185087	AC007475	Drosophila melanogaster chromosome 2 clone BACR04E21 (D592) RPCI-98 04.E.21 map 49A-49B strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***, 36 unordered pieces.	Drosophila melanogaster	37,333	2-Aug-99
rx01776	1575	GB_GSS10: AQ174954	416	AQ174954	HS_3211_B2_C03_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plates=3211 Col=6 Row=F, genomic survey sequence.	Homo sapiens	43,171	17-OCT-1998
		GB_PR1: HSGFG6	1032	X57075	H. sapiens FGF6 gene.	Homo sapiens	41,196	15-Feb-95
		GB_PR2: HSAC000055	93578	AC000055	Human PAC clone DJ073F11 from Xq23, complete sequence.	Homo sapiens	35,080	14-Nov-96
rx01777	582	GB_BA2: AF076997	15398	AF076997	Comamonas testosteroni PLS5 cryptic plasmid pPT1, complete sequence.	Comamonas testosteroni	37,805	30-Nov-99
		GB_BA2: AF076997	15398	AF076997	Comamonas testosteroni PLS5 cryptic plasmid pPT1, complete sequence.	Comamonas testosteroni	35,529	30-Nov-99
rx01778	1260	GB_PR3: AC002981	153588	AC002981	Homo sapiens Xp22 BAC GS279A12 (Genome Systems) complete sequence.	Homo sapiens	37,305	22-Jan-98
		GB_EST29: AI587176	384	AI587176	tr54c10.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2222130 3', mRNA sequence.	Homo sapiens	37,240	14-MAY-1999
		GB_PR1: HSCP450	1346	X65962	H. sapiens mRNA for cytochrome P-450.	Homo sapiens	38,908	29-MAY-1992
rx01779	1542	GB_EST19: AA202518	661	AA202518	LD20757 5prime LD Drosophila melanogaster embryo BlueScript Drosophila melanogaster cDNA clone LD02757 5prime, mRNA sequence.	Drosophila melanogaster	40,212	27-Nov-98
		GB_EST34: AI805834	453	AI805834	te52b03.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2090285 3' similar to contains Alu repetitive element; mRNA sequence.	Homo sapiens	38,000	7-Jul-99
		GB_EST30: AI648920	754	AI648920	uk32h03.x1 Sugano mouse kidney mRna Mus musculus cDNA clone IMAGE:1970741 3' similar to WP:Y76A2B.5 CE19277; mRNA sequence.	Mus musculus	36,967	30-Apr-99
rx01780	498	GB_GSS8: AQ038759	670	AQ038759	CIT-HSP-2325O13.TV CIT-HSP Homo sapiens genomic clone 2325O13, genomic survey sequence.	Homo sapiens	32,645	11-Jul-98
		GB_GSS8: AQ038759	670	AQ038759	CIT-HSP-2325O13.TV CIT-HSP Homo sapiens genomic clone 2325O13, genomic survey sequence.	Homo sapiens	39,474	11-Jul-98
rx01781	384	GB_IN1: CELF52G3	42696	AF026212	Caenorhabditis elegans cosmid F52G3.	Caenorhabditis elegans	39,314	25-Sep-97
		GB_GSS14: AQ579313	877	AQ579313	nbxb0084O07r: CUGI Rice BAC Library Oryza sativa genomic clone nbxb0084O07r, genomic survey sequence.	Oryza sativa	36,053	2-Jun-99
		GB_VI: HAV19AE3	3441	X95259	Human adenovirus type 19a early region 3 genes.	Human adenovirus type 19a	38,624	9-Jul-97
rx01782	519	GB_IN1: DMS5ED5	930	X78219	D. melanogaster (Barton) SED5 mRNA.	Drosophila melanogaster	38,372	21-Apr-95
		GB_IN2: L49408	83527	L49408	Drosophila melanogaster DNA sequence (P1 DS02740 (D27)), complete sequence.	Drosophila melanogaster	35,992	17-Jul-98
		GB_STS: G46002	398	G46002	Z556_1 Zebrafish AB Danio rerio STS genomic clone Z556 5', sequence tagged site.	Danio rerio	33,920	23-MAR-1999
rx01783	360	GB_VI: CITV18420	19259	Y18420	Citrus tristeza virus complete genome, isolate T385.	Citrus tristeza virus	41,457	6-Apr-99
		GB_IN1: DMS5ED5	930	X78219	D. melanogaster (Barton) SED5 mRNA.	Drosophila melanogaster	45,802	21-Apr-95
		GB_VI: CTV18368	2759	Y18368	Citrus tristeza virus defective RNA, strain T411.	Citrus tristeza virus	41,457	6-Apr-99
rx01785	699	GB_BA2: AF121000	19751	AF121000	Corynebacterium glutamicum strain 22243 R-plasmid pAG1, complete sequence.	Corynebacterium glutamicum	40,529	14-Apr-99
		GB_PR2: HS1014D13	71263	AL022311	Human DNA sequence from clone 1014D13 on chromosome 22q13.1-13.2 Contains ESTs, STSs, and a CpG island, complete sequence.	Homo sapiens	35,362	23-Nov-99
		GB_BA2: AF121000	19751	AF121000	Corynebacterium glutamicum strain 22243 R-plasmid pAG1, complete sequence.	Corynebacterium glutamicum	38,012	14-Apr-99
rx01787	816	GB_RO: RATSIALYLA	1128	L13445	Rat sialyltransferase gene family related mRNA, complete cds.	Rattus norvegicus	35,012	26-Jul-93
		GB_PR3: HSN75H12	46144	Z84496	Human DNA sequence from cosmid N75H12 on chromosome 22q12-qter.	Homo sapiens	37,738	23-Nov-99



**TABLE 4: ALIGNMENT RESULTS**

GB_GSS13: AQ496514	532	AQ496514	HS_5118_B2_F05_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=694 Col=10 Row=L, genomic survey sequence.	Homo sapiens	39,238	28-Apr-99
GB_BA1: MTV029	3279	AL021427	Mycobacterium tuberculosis H37Rv complete genome; segment 1/162.	Mycobacterium tuberculosis	39,229	17-Jun-98
GB_BA1: MTORIREP	8400	X92504	Mt.tuberculosis origin of replication and genes mpA, rpmH, dnaA, dnaN, recF.	Mycobacterium tuberculosis	36,328	26-Aug-97
GB_GSS4: AQ729452	763	AQ729452	HS_5474_A1_G07_TTAHS_5474_A1 RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1050 Col=13 Row=M, genomic survey sequence.	Homo sapiens	39,451	15-Jul-99
GB_EST37: A1935426	436	A1935426	w084002.x1 NCL_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2462019 3' similar to TR:O00193 O00193 SMALL ACIDIC PROTEIN, mRNA sequence.	Homo sapiens	39,877	2-Sep-99
GB_GSS11: AQ265301	621	AQ265301	CITBLE1-2509K16.1F CITBLE1 Homo sapiens genomic clone 2509K16, genomic survey sequence.	Homo sapiens	45,856	27-OCT-1998
GB_HTG2: AC007596	199300	AC007596	Homo sapiens chromosome 16 clone 11686, *** SEQUENCING IN PROGRESS *** 42 unordered pieces.	Homo sapiens	37,572	20-MAY-1999
GB_PL12: ATAC004521	104797	AC004521	Arabidopsis thaliana chromosome II BAC F411 genomic sequence, complete sequence.	Arabidopsis thaliana	43,167	12-MAY-1998
GB_GSS13: AQ436125	508	AQ436125	HS_5049_B2_G12_TTA RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=625 Col=24 Row=N, genomic survey sequence.	Homo sapiens	32,762	31-MAR-1999
GB_GSS12: AQ342410	674	AQ342410	RPCI11-121E5.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-121E5, genomic survey sequence.	Homo sapiens	37,269	06-MAY-1999
GB_STS: KLAJ9905	178	AJ229905	Kluyveromyces lactis DNA fragment for sequence tagged site, clone okam5e06r.	Kluyveromyces lactis	42,697	20-Nov-98
GB_STS: KLAJ9905	178	AJ229905	Kluyveromyces lactis DNA fragment for sequence tagged site, clone okam5e06r.	Kluyveromyces lactis	42,697	20-Nov-98
GB_GSS15: AQ612295	472	AQ612295	HS_5121_A2_D09_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=697 Col=18 Row=G, genomic survey sequence.	Homo sapiens	40,885	15-Jun-99
GB_GSS3: B82515	613	B82515	RPCI11-16J10.TV RPCI-11 Homo sapiens genomic clone RPCI-11-16J10, genomic survey sequence.	Homo sapiens	40,164	9-Apr-99
GB_PR2: CNS0000B	196287	AL049829	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-124D2 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.	Homo sapiens	33,621	27-OCT-1999
GB_OV: AF001393	4329	AF001393	Oryzias latipes Medaka OG-12 (MOG-12) mRNA, complete cds.	Oryzias latipes	35,511	30-Sep-99
GB_GSS15: AQ644157	665	AQ644157	RPCI93-Dpnll-29O12.TV RPCI93-Dpnll Trypanosoma brucei genomic clone RPCI93-Dpnll-29O12, genomic survey sequence.	Trypanosoma brucei	39,894	8-Jul-99
GB_GSS15: AQ657704	665	AQ657704	Sheared DNA Trypanosoma brucei genomic clone Sheared DNA-24C24, genomic survey sequence.	Trypanosoma brucei	37,110	23-Jun-99
GB_OV: GGU25125	6418	U25125	Gallus gallus preprogastrin gene, complete cds.	Gallus gallus	40,584	06-MAY-1995
GB_PR3: HS404G5	119737	AL035695	Human DNA sequence from clone 404G5 on chromosome 6q24.1-25.2. Contains part of a human estrogen receptor gene, STSs and GSSs, complete sequence.	Homo sapiens	34,793	23-Nov-99
GB_PR1: HSPPTP:CHG	8545	X82818	H sapiens PTP1C/HCP gene.	Homo sapiens	35,531	26-Jun-97
GB_PR3: HS404G5	119737	AL035695	Human DNA sequence from clone 404G5 on chromosome 6q24.1-25.2. Contains part of a human estrogen receptor gene, STSs and GSSs, complete sequence.	Homo sapiens	37,058	23-Nov-99
GB_HTG3: AC010869	38000	AC010869	Leishmania major chromosome 35 clone L7195 strain Friedlin, *** SEQUENCING IN PROGRESS *** 4 unordered pieces.	Leishmania major	40,522	02-OCT-1999
GB_HTG3: AC010869	38000	AC010869	Leishmania major chromosome 35 clone L7195 strain Friedlin, *** SEQUENCING IN PROGRESS *** 4 unordered pieces.	Leishmania major	40,522	02-OCT-1999
GB_OM: SSU12574	2190	U12574	Sus scrofa myogenic regulatory factor MyoD (myoD) gene, complete cds.	Sus scrofa	37,068	10-Feb-96
GB_PH: C2PVC	22172	L48605	Bacteriophage c2 complete genome.	Lactococcus bacteriophage c2	36,747	14-MAR-1996
GB_EST38: AW066009	641	AW066009	687004F08.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA sequence.	Zea mays	40,000	12-OCT-1999
GB_PH: C2PVC	22172	L48605	Bacteriophage c2 complete genome.	Lactococcus bacteriophage c2	36,867	14-MAR-1996

TABLE 4: ALIGNMENT RESULTS

rx01803	1083	GB_BA2:SCJ1	36925	AL109962	Streptomyces coelicolor cosmid J1.	Streptomyces coelicolor A3(2)	54,267	24-Sep-99
		GB_BA1:PAU12891	4062	U12891	Pseudomonas aeruginosa PAO substrain OT684 pyoverdine gene transcriptional regulator PvdS (pvdS) gene, complete cds.	Pseudomonas aeruginosa	37,841	03-OCT-1996
		GB_HTG6:AC010203	230460	AC010203	Homo sapiens clone RP11-175P13, *** SEQUENCING IN PROGRESS ***, 48 unordered pieces.	Homo sapiens	35,185	03-DEC-1999
rx01804	444	GB_EST24:AI189912	784	AI189912	qd33e07.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA clone IMAGE:1725540 3' similar to gb:Z23064_cds1 HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN G (HUMAN); mRNA sequence.	Homo sapiens	40,000	28-OCT-1998
		GB_EST24:AI189912	784	AI189912	qd33e07.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA clone IMAGE:1725540 3' similar to gb:Z23064_cds1 HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN G (HUMAN); mRNA sequence.	Homo sapiens	35,294	28-OCT-1998
rx01805	453	GB_PR2:HSU73628	32289	U73628	Human chromosome 11 101h11 cosmid, complete sequence.	Homo sapiens	36,099	19-Jun-97
		GB_PR2:HSU73628	32289	U73628	Human chromosome 11 101h11 cosmid, complete sequence.	Homo sapiens	35,227	19-Jun-97
rx01806	603	GB_IN1:CELK02E7	38535	AF025465	Caenorhabditis elegans cosmid K02E7.	Caenorhabditis elegans	37,479	23-Sep-97
		GB_HTG3:AC011217	157261	AC011217	Homo sapiens clone 7_J_14, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	33,333	03-OCT-1999
		GB_HTG3:AC011217	157261	AC011217	Homo sapiens clone 7_J_14, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	33,333	03-OCT-1999
rx01809	1299	GB_HTG3:AC010759	155875	AC010759	Homo sapiens clone 1_K_15, *** SEQUENCING IN PROGRESS ***, 15 unordered pieces.	Homo sapiens	37,920	22-Sep-99
		GB_HTG3:AC010759	155875	AC010759	Homo sapiens clone 1_K_15, *** SEQUENCING IN PROGRESS ***, 15 unordered pieces.	Homo sapiens	37,920	22-Sep-99
		GB_HTG2:AC006235	190842	AC006235	Homo sapiens clone hRPK.520_J_4, *** SEQUENCING IN PROGRESS ***, 5 unordered pieces.	Homo sapiens	36,019	17-Jul-99
rx01813	789	GB_PR4:AC004047	134649	AC004047	Homo sapiens chromosome 4 clone B150J.4 map 4q25, complete sequence.	Homo sapiens	36,205	01-DEC-1998
		GB_HTG2:HS1012F16	74539	AL080274	Homo sapiens chromosome 20 clone RP5-1012F16, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Homo sapiens	35,032	03-DEC-1999
		GB_HTG2:HS1012F16	74539	AL080274	Homo sapiens chromosome 20 clone RP5-1012F16, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Homo sapiens	35,032	03-DEC-1999
rx01815	915	GB_PR3:AC005740	186780	AC005740	Homo sapiens chromosome 5p, BAC clone 50g21 (LBNL H154), complete sequence.	Homo sapiens	37,315	01-OCT-1998
		GB_HTG3:AC008328	114617	AC008328	Drosophila melanogaster chromosome 2 clone BACR09A04 (D860) RPCI-98 09 A.4 map 28B-28C strain Y, cn bw sp, *** SEQUENCING IN PROGRESS ***, 74 unordered pieces.	Drosophila melanogaster	30,516	6-Aug-99
		GB_HTG3:AC008328	114617	AC008328	Drosophila melanogaster chromosome 2 clone BACR09A04 (D860) RPCI-98 09 A.4 map 28B-28C strain Y, cn bw sp, *** SEQUENCING IN PROGRESS ***, 74 unordered pieces.	Drosophila melanogaster	30,516	6-Aug-99
rx01816	855	GB_PR3:AC003101	208396	AC003101	Homo sapiens chromosome 17, clone HRP41C23, complete sequence.	Homo sapiens	34,272	5-Jun-98
		GB_EST23:AI144656	492	AI144656	UI-R-BT0-pl-f-04-0-UI.s1 UI-R-BT0 Rattus norvegicus cDNA clone UI-R-BT0-pl-f-04-0-UI 3', mRNA sequence.	Rattus norvegicus	37,526	5-Jul-99
		GB_GSS11:AQ341706	506	AQ341706	RPCI11-120J20.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-120J20, genomic survey sequence.	Homo sapiens	40,789	06-MAY-1999
rx01817	756	GB_BA2:AF030176	7000	AF030176	Rhodococcus opacus putative transposase gene, partial cds; and putative FAD synthetase, putative short-chain dehydrogenase/reductase, maleylacetate reductase (macA), and putative transcription factor genes, complete cds.	Rhodococcus opacus	42,577	5-Aug-98
		GB_EST19:AA263350	423	AA263350	LD06495.5prime LD Drosophila melanogaster embryo BlueScript Drosophila melanogaster cDNA clone LD06495 5prime, mRNA sequence.	Drosophila melanogaster	38,955	27-Nov-98
		GB_EST19:AA202194	368	AA202194	LD02304.5prime LD Drosophila melanogaster embryo BlueScript Drosophila melanogaster cDNA clone LD02304 5prime, mRNA sequence.	Drosophila melanogaster	39,891	27-Nov-98
rx01820	1497	GB_BA2:AF030176	7000	AF030176	Rhodococcus opacus putative transposase gene, partial cds; and putative FAD synthetase, putative short-chain dehydrogenase/reductase, maleylacetate reductase (macA), and putative transcription factor genes, complete cds.	Rhodococcus opacus	41,799	5-Aug-98
		GB_EST32:AI728582	666	AI728582	BNLGH1124 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U64918) ATGP1 [Arabidopsis thaliana], mRNA sequence.	Gossypium hirsutum	37,915	11-Jun-99

TABLE 4: ALIGNMENT RESULTS

rx01825	393	GB_BA2:AF030176	7000	AF030176	Rhodococcus opacus putative transposase gene, partial cds; and putative FAD synthetase, putative short-chain dehydrogenase/reductase, maleylacetate reductase (macA), and putative transcription factor genes, complete cds.	Rhodococcus opacus	37,127	5-Aug-98
		GB_HTG1:HSJCF13	293368	AJ239320	Homo sapiens chromosome X clone cosmid 244B12 map Xq13, *** SEQUENCING IN PROGRESS	Homo sapiens	34,748	28-Sep-99
		GB_HTG1:HSJCF13	293368	AJ239320	*** in ordered pieces.			
		GB_HTG1:AP000568	136627	AP000568	Homo sapiens chromosome X clone cosmid 244B12 map Xq13, *** SEQUENCING IN PROGRESS	Homo sapiens	34,748	28-Sep-99
		GB_IN2:AF047757	408	AF047757	Homo sapiens chromosome 21 clone B753B2 map 21q21.2, *** SEQUENCING IN PROGRESS	Homo sapiens	34,884	07-OCT-1999
rx01831	585	GB_EST18:AA687516	513	AA687516	in unordered pieces.			
		GB_EST18:AA687516	513	AA687516	Trichogramma australicum internal transcribed spacer 2, complete sequence.	Trichogramma australicum	38,390	10-Jul-98
		GB_EST18:AA687516	513	AA687516	ns58h11.1 NCL_CGAP_P122 Homo sapiens cDNA clone IMAGE:1187877 5' similar to gb:X15341 CYTOCHROME C OXIDASE POLYPEPTIDE VIA-LIVER (HUMAN), mRNA sequence.	Homo sapiens	35,057	11-DEC-1997
		GB_EST18:AA687516	513	AA687516	ns58h11.1 NCL_CGAP_P122 Homo sapiens cDNA clone IMAGE:1187877 5' similar to gb:X15341 CYTOCHROME C OXIDASE POLYPEPTIDE VIA-LIVER (HUMAN), mRNA sequence.	Homo sapiens	37,624	11-DEC-1997
rx01834	825	GB_PL2:T8F5	87192	AC004512	Arabidopsis thaliana chromosome 1 BAC T8F5 sequence, complete sequence.	Arabidopsis thaliana	38,101	22-Jul-98
		GB_PR4:AC007677	166949	AC007677	Homo sapiens clone NH0086N01, complete sequence.	Homo sapiens	35,888	28-Sep-99
		GB_PR4:AC007677	166949	AC007677	Homo sapiens clone NH0086N01, complete sequence.	Homo sapiens	36,906	28-Sep-99
rx01842	1041	GB_HTG3:AC009183	145694	AC009183	Drosophila melanogaster chromosome 3 clone BACR39F04 (D839) RPCI-98 39.F.4 map 85F-85F strain Y; cn bw sp. *** SEQUENCING IN PROGRESS	Drosophila melanogaster	40,430	16-Sep-99
		GB_HTG3:AC009183	145694	AC009183	*** SEQUENCING IN PROGRESS			
		GB_HTG5:AC008338	136685	AC008338	Drosophila melanogaster chromosome 3 clone BACR39F04 (D839) RPCI-98 39.F.4 map 85F-85F strain Y; cn bw sp. *** SEQUENCING IN PROGRESS	Drosophila melanogaster	40,430	16-Sep-99
		GB_HTG5:AC008338	136685	AC008338	*** SEQUENCING IN PROGRESS			
rx01844	303	GB_HTG2:AC007597	163880	AC007597	Drosophila melanogaster chromosome X clone BACR30J04 (D908) RPCI-98 30.J.4 map 19C-19E strain Y; cn bw sp. *** SEQUENCING IN PROGRESS	Drosophila melanogaster	40,527	15-Nov-99
		GB_HTG2:AC007597	163880	AC007597	*** SEQUENCING IN PROGRESS			
		GB_HTG3:AC009998	163590	AC009998	Homo sapiens chromosome 16 clone RPCI-11_137H10, *** SEQUENCING IN PROGRESS	Homo sapiens	34,564	20-Jul-99
		GB_HTG3:AC009998	163590	AC009998	ordered pieces.			
		GB_OV:AF012348	4921	AF012348	Homo sapiens chromosome 16 clone RPCI-11_137H10, *** SEQUENCING IN PROGRESS	Homo sapiens	34,564	20-Jul-99
rx01845	531	GB_BA1:D85417	7984	D85417	ordered pieces.			
		GB_BA1:D85417	7984	D85417	Homo sapiens chromosome 15 clone BAC 573G7 map 15q24, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	40,816	10-Sep-99
		GB_OV:AF012348	4921	AF012348	Gallus gallus smooth muscle gamma actin (gamma actin) gene, complete cds.	Gallus gallus	37,333	17-MAR-1998
		GB_BA1:D85417	7984	D85417	Propionibacterium freudenreichii hemY, hemH, hemB, hemX, hemR and hemL genes, complete cds.	Propionibacterium freudenreichii	45,243	6-Feb-99
		GB_BA1:D85417	7984	D85417	Propionibacterium freudenreichii hemY, hemH, hemB, hemX, hemR and hemL genes, complete cds.	Propionibacterium freudenreichii	40,232	6-Feb-99
rx01846	382	GB_HTG1:HSJ323A24	278948	AL121750	Propionibacterium freudenreichii hemY, hemH, hemB, hemX, hemR and hemL genes, complete cds.	Propionibacterium freudenreichii	34,840	23-Nov-99
		GB_HTG1:HSJ323A24	278948	AL121750	Homo sapiens chromosome 4 clone RP3-323A24, *** SEQUENCING IN PROGRESS	Homo sapiens	34,840	23-Nov-99
		GB_HTG1:HSJ323A24	278948	AL121750	ordered pieces.			
		GB_PR2:HS181C9	92472	Z98743	Homo sapiens chromosome 4 clone RP3-323A24, *** SEQUENCING IN PROGRESS	Homo sapiens	42,480	23-Nov-99
		GB_PR2:HS181C9	92472	Z98743	ordered pieces.			
rx01847	858	GB_PR3:AC004893	103738	AC004893	Human DNA sequence from clone 181C9 on chromosome 22q13.2-13.33 Contains a PHAPI2 Leucine Rich Acidic Nuclear Protein pseudogene, part of a putative novel gene, ESTs, STSs and GSSs, complete sequence.	Homo sapiens	38,321	03-OCT-1998
		GB_HTG4:AC010182_3	110000	AC010182	Homo sapiens PAC clone DJ0808A01 from 7q21.1-q31.1, complete sequence.	Homo sapiens	36,998	15-Sep-99
		GB_HTG4:AC010182_3	110000	AC010182	Homo sapiens chromosome 12q seeders clone RPCI11-185N2, *** SEQUENCING IN PROGRESS	Homo sapiens		
		GB_HTG4:AC010182_3	110000	AC010182	***, 172 unordered pieces.			
		GB_HTG4:AC010182_3	110000	AC010182	Homo sapiens chromosome 12q seeders clone RPCI11-185N2, *** SEQUENCING IN PROGRESS	Homo sapiens	36,998	15-Sep-99
		GB_HTG4:AC010182_3	110000	AC010182	***, 172 unordered pieces.			
rx01856	669	GB_GSS15:AQ612487	433	AQ612487	HS_5122_A1_B07_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=698 Col=13 Row=C, genomic survey sequence.	Homo sapiens	39,723	15-Jun-99

TABLE 4: ALIGNMENT RESULTS

rx01857	774	GB_PR4:AC005876	89817	AC005876	ctb_188_i_5, complete sequence.	Homo sapiens	39,567	5-Nov-99
		GB_PR4:AC005876	89817	AC005876	ctb_188_i_5, complete sequence.	Homo sapiens	36,677	5-Nov-99
		GB_PR3:HS742C19	122748	AL031846	Human DNA sequence from clone RP4-742C19 on chromosome 22, complete sequence.	Homo sapiens	33,506	29-Nov-99
		GB_PR2:HSJ110016	161525	AL110503	Human DNA sequence from clone RP5-110016 on chromosome 20. Contains a novel mRNA, GSSs and a CpG island, complete sequence.	Homo sapiens	37,450	22-Nov-99
		GB_HTG6:AC009893	186769	AC009893	Homo sapiens chromosome 8 clone RP11-4P3, *** SEQUENCING IN PROGRESS ***; 5 unordered pieces.	Homo sapiens	37,368	23-Nov-99
rx01858								
rx01870	798	GB_RO:AC004155	128026	AC004155	Mus musculus DNA from BAC 10818 containing the Ercc-4 gene, complete sequence.	Mus musculus	37,387	19-Feb-98
		GB_RO:AC004155	128026	AC004155	Mus musculus DNA from BAC 10818 containing the Ercc-4 gene, complete sequence.	Mus musculus	37,707	19-Feb-98
		GB_BA2:AF144563	5971	AF144563	Thermobifida fusca beta-1,4-exocellulase E6 precursor (celf) gene, complete cds, and unknown genes.	Thermobifida fusca	38,619	21-Jun-99
rx01871	1086	GB_BA2:AF079304	3350	AF079304	Eikenella corrodens type IV pilin (pilA1), type IV pilin (pilA2), putative fibrin protein (pilB), and putative hemagglutinin protein (hagA) genes, complete cds.	Eikenella corrodens	40,465	15-Jul-99
		GB_BA1:ECECPHAG	3589	Z12609	E. corrodens ecpA and ecpB genes encoding type 4 N-methylphenylalanine pilin and hag1 gene for hemagglutinin protein.	Eikenella corrodens	40,517	5-Aug-92
rx01874	448	GB_BA2:AE001707	19518	AE001707	Thermotoga maritima section 19 of 136 of the complete genome.	Thermotoga maritima	38,124	2-Jun-99
		GB_EST6:N51407	505	N51407	yz17a08.s1 Soares_multiple_sclerosis_2NbHMSF Homo sapiens cDNA clone IMAGE:283286 3', mRNA sequence.	Homo sapiens	35,735	14-Feb-96
		GB_EST29:A1619782	374	A1619782	yz17a08.s1 Soares_multiple_sclerosis_2NbHMSF Homo sapiens cDNA clone IMAGE:283286 3', mRNA sequence.	Homo sapiens	38,966	21-Apr-99
		GB_EST36:A1885047	466	A1885047	w189a01.x1 NC1_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2432040 3', mRNA sequence.	Homo sapiens	37,419	1-Sep-99
rx01875	969	GB_EST13:AA356276	291	AA356276	EST84843 Jurkat T-cells VI Homo sapiens cDNA 5' end, mRNA sequence.	Homo sapiens	41,667	21-Apr-97
		GB_EST22:A1050653	443	A1050653	ub38f03.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1380029 5', mRNA sequence.	Mus musculus	38,851	9-Jul-98
		GB_GSS1:CNS0022Z	1101	AL061786	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR05B21 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	Drosophila melanogaster	39,039	3-Jun-99
rx01877	1194	GB_HTG4:AC009849	114993	AC009849	Drosophila melanogaster chromosome 2 clone BACR07H08 (D864) RPCI-98 07 H.8 map 31B-31C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 55 unordered pieces.	Drosophila melanogaster	37,274	25-OCT-1999
		GB_HTG4:AC009849	114993	AC009849	Drosophila melanogaster chromosome 2 clone BACR07H08 (D864) RPCI-98 07 H.8 map 31B-31C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 55 unordered pieces.	Drosophila melanogaster	37,274	25-OCT-1999
		GB_IN2:AC005454	84367	AC005454	Drosophila melanogaster, chromosome 2R, region 31C1-31D6, P1 clone DS08879, complete sequence.	Drosophila melanogaster	38,679	15-DEC-1998
rx01879	1056	GB_PR2:CNS0000Q	192932	AL049874	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-1042B17 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.	Homo sapiens	37,094	7-Jul-99
		GB_HTG2:AC006732	159453	AC006732	Caenorhabditis elegans clone Y32G9, *** SEQUENCING IN PROGRESS ***; 9 unordered pieces.	Caenorhabditis elegans	34,862	23-Feb-99
		GB_HTG2:AC006732	159453	AC006732	Caenorhabditis elegans clone Y32G9, *** SEQUENCING IN PROGRESS ***; 9 unordered pieces.	Caenorhabditis elegans	34,862	23-Feb-99
rx01896	759	GB_BA1:SC2G5	38404	AL035478	Streptomyces coelicolor cosmid 2G5.	Streptomyces coelicolor	38,420	11-Jun-99
		GB_PL1:AP000570	157903	AP000570	Oryza sativa genomic DNA, chromosome 1, clone:P0711E10.	Oryza sativa	40,272	03-DEC-1999
		GB_BA1:SC2G5	38404	AL035478	Streptomyces coelicolor cosmid 2G5.	Streptomyces coelicolor	40,270	11-Jun-99
rx01899	909	GB_BA1:MTV002	56414	AL008967	Mycobacterium tuberculosis H37Rv complete genome, segment 122/162.	Mycobacterium tuberculosis	38,702	17-Jun-98
		GB_BA1:MTV002	56414	AL008967	Mycobacterium tuberculosis H37Rv complete genome, segment 122/162.	Mycobacterium tuberculosis	38,229	17-Jun-98
rx01902	1182	GB_HTG6:AC009893	186769	AC009893	Homo sapiens chromosome 8 clone RP11-4P3, *** SEQUENCING IN PROGRESS ***; 5 unordered pieces.	Homo sapiens	34,868	23-Nov-99
		GB_OV:D84063	2363	D84063	Oryzias latipes Bf/C2 mRNA, complete cds.	Oryzias latipes	38,074	6-Feb-99

**TABLE 4: ALIGNMENT RESULTS**

rx01903	302	GB_HTG6:AC009893	186769	AC009893	Homo sapiens chromosome 8 clone RP11-4P3, *** SEQUENCING IN PROGRESS ***	5 unordered	Homo sapiens	35,616	23-Nov-99
		GB_BA1:MSGILVB	4210	L49392	Myobacterium avium acetolactate synthase (ilvBN) and acetohydroxy acid isomeroreductase (ilvC) gene, complete cds.		Myobacterium avium	38,667	11-DEC-1996
		GB_PR3:AC004833	68890	AC004833	Homo sapiens PAC clone DJ0547C10 from 7p21-p22, complete sequence.		Homo sapiens	34,615	5-Nov-98
		GB_PR3:AC004833	68890	AC004833	Homo sapiens PAC clone DJ0547C10 from 7p21-p22, complete sequence.		Homo sapiens	42,712	5-Nov-98
rx01904	546	GB_HTG2:AC006740	200965	AC006740	Caenorhabditis elegans clone Y38B5, *** SEQUENCING IN PROGRESS ***	12 unordered	Caenorhabditis elegans	34,074	25-Feb-99
		GB_HTG2:AC006695	33622	AC006695	Caenorhabditis elegans clone W06H8, *** SEQUENCING IN PROGRESS ***	1 unordered	Caenorhabditis elegans	35,556	23-Feb-99
		GB_HTG2:AC006695	33622	AC006695	Caenorhabditis elegans clone W06H8, *** SEQUENCING IN PROGRESS ***	1 unordered	Caenorhabditis elegans	35,556	23-Feb-99
rx01905	654	GB_IN1:CEW04G3	32158	Z68014	Caenorhabditis elegans cosmid W04G3, complete sequence.		Caenorhabditis elegans	38,730	2-Sep-99
		GB_PL2:AF067082	6947	AF067082	Apium graveolens mannitol dehydrogenase (Mtd) gene, complete cds.		Apium graveolens	38,242	30-Sep-98
		GB_IN1:CEW04G3	32158	Z68014	Caenorhabditis elegans cosmid W04G3, complete sequence.		Caenorhabditis elegans	36,834	2-Sep-99
rx01906	588	GB_PL1:AB012242	78973	AB012242	Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone: K24G6, complete sequence.		Arabidopsis thaliana	39,071	20-Nov-99
		GB_EST6:N49608	454	N49608	yy58g01.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone IMAGE:277776 3', mRNA sequence.		Homo sapiens	43,603	14-Feb-96
		GB_EST6:N49609	452	N49609	yy58g02.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone IMAGE:277778 3', mRNA sequence.		Homo sapiens	43,307	14-Feb-96
rx01907	735	GB_HTG3:AC009437	159691	AC009437	Homo sapiens chromosome 11 clone 56_G_09 map 11, *** SEQUENCING IN PROGRESS ***	10 unordered	Homo sapiens	36,226	22-Aug-99
		GB_HTG3:AC009437	159691	AC009437	Homo sapiens chromosome 11 clone 56_G_09 map 11, *** SEQUENCING IN PROGRESS ***	10 unordered	Homo sapiens	36,226	22-Aug-99
		GB_HTG3:AC009437	159691	AC009437	Homo sapiens chromosome 11 clone 56_G_09 map 11, *** SEQUENCING IN PROGRESS ***	10 unordered	Homo sapiens	38,483	22-Aug-99
rx01908	969	GB_PR3:HS90G24	154414	AL008723	Human DNA sequence from clone 90G24 on chromosome 22. Contains the RPPL2 gene for RET finger protein-like 2, an immunoglobulin Lambda Light Chain C region (IGLC) pseudogene, the gene for SAAT1 (low affinity sodium glucose cotransporter (sodium:solute symporter family)) and a Cleavage and Polyadenylation Specific Factor CPSF 160 kD subunit pseudogene. Contains ESTs, GSSs and three putative CpG islands, complete sequence.		Homo sapiens	37,331	23-Nov-99
		GB_EST25:AI316482	972	AI316482	uj60g12.y1 Sugano mouse liver mRna Mus musculus cDNA clone IMAGE:1924390 5' similar to gb:M12529 APOLIPROTEIN E PRECURSOR (HUMAN); gb:D00466 Mouse apolipoprotein E gene (MOUSE), mRNA sequence.		Mus musculus	36,800	17-DEC-1998
rx01909	970	GB_EST27:AI398904	556	AI398904	NCW10G777 Westergaards Neurospora crassa cDNA clone W10G7 3', mRNA sequence.		Neurospora crassa	38,095	8-Feb-99
		GB_HTG1:HSJ65814	133423	AL109845	Homo sapiens chromosome 1 clone RP4-65814, *** SEQUENCING IN PROGRESS ***	in unordered	Homo sapiens	40,520	23-Nov-99
		GB_HTG1:HSJ65814	133423	AL109845	Homo sapiens chromosome 1 clone RP4-65814, *** SEQUENCING IN PROGRESS ***	in unordered	Homo sapiens	40,520	23-Nov-99
		GB_HTG1:HSJ65814	133423	AL109845	Homo sapiens chromosome 1 clone RP4-65814, *** SEQUENCING IN PROGRESS ***	in unordered	Homo sapiens	36,364	23-Nov-99
rx01910	969	GB_PR2:CNS00008	218956	AL049831	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-330O19 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.		Homo sapiens	35,393	29-Jun-99
		GB_PR2:CNS00008	218956	AL049831	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-330O19 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.		Homo sapiens	38,405	29-Jun-99
		GB_HTG5:AC009217	118561	AC009217	Drosophila melanogaster chromosome X clone BACR41N19 (D907) RPCI:98 41.N.19 map 19A-19C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***	78 unordered	Drosophila melanogaster	36,765	16-Nov-99
rx01911	1062	GB_HTG2:AC002317	94882	AC002317	Homo sapiens chromosome 17 clone HClT7H10 map 17, *** SEQUENCING IN PROGRESS ***	8 unordered	Homo sapiens	36,654	20-Feb-98
		GB_HTG2:AC002317	94882	AC002317	Homo sapiens chromosome 17 clone HClT7H10 map 17, *** SEQUENCING IN PROGRESS ***	8 unordered	Homo sapiens	36,654	20-Feb-98

TABLE 4: ALIGNMENT RESULTS

rx01923	873	GB_HTG2:AC002317	94882	AC002317	Homo sapiens chromosome 17 clone HC17H10 map 17, *** SEQUENCING IN PROGRESS ***	8	Homo sapiens	37,012	20-Feb-98
		GB_HTG2:AC006844	298782	AC006844	unordered pieces.		Caenorhabditis elegans	37,176	24-Feb-99
		GB_HTG2:AC006844	298782	AC006844	Caenorhabditis elegans clone Y108G3Y, *** SEQUENCING IN PROGRESS ***	4	Caenorhabditis elegans	37,176	24-Feb-99
		GB_BA1:ECOUW76	225419	U00039	Caenorhabditis elegans clone Y108G3Y, *** SEQUENCING IN PROGRESS ***	4	Caenorhabditis elegans	37,176	24-Feb-99
rx01930	1074	GB_BA1:CGPAN	2164	X96580	pieces.		Escherichia coli	38,902	7-Nov-96
		GB_HTG2:AC007598	248427	AC007598	E. coli chromosomal region from 76.0 to 81.5 minutes.		Corynebacterium glutamicum	100,000	11-MAY-1999
		GB_HTG2:AC007598	248427	AC007598	C. glutamicum panB, panC & xylB genes.		Homo sapiens	38,469	20-MAY-1999
		GB_IN1:CEH12119	37427	Z98851	Homo sapiens chromosome 16 clone 165M1, *** SEQUENCING IN PROGRESS ***	105	Homo sapiens	38,469	20-MAY-1999
rx01931	777	GB_HTG1:CEY37A1	316170	Z93243	unordered pieces.		Caenorhabditis elegans	32,763	18-DEC-1998
		GB_HTG1:CEY37A1	316170	Z93243	Caenorhabditis elegans cosmid H12119, complete sequence.		Caenorhabditis elegans	32,763	03-DEC-1998
rx01941	558	GB_EST25:A1313927	520	A1313927	Caenorhabditis elegans chromosome IV clone Y37A1, *** SEQUENCING IN PROGRESS ***	in	Caenorhabditis elegans	32,763	03-DEC-1998
		GB_RO:AF053757	8308	AF053757	unordered pieces.		Mus musculus	41,148	17-DEC-1998
rx01942	723	GB_RO:MMU77461	2657	U77461	Caenorhabditis elegans chromosome IV clone Y37A1, *** SEQUENCING IN PROGRESS ***	in	Mus musculus	38,649	30-Jul-98
		GB_BA1:BRLPTSG	3163	L18875	TR:O09047 O09047 COMPLEMENT COMPONENT 3A RECEPTOR 1, mRNA sequence.		Mus musculus	38,649	28-MAY-1997
		GB_BA1:CORPTSMA	2656	L18874	Mus musculus anaphylatoxin C3a receptor gene, complete cds.		Brevibacterium lactofermentum	67,407	01-OCT-1993
		GB_BA1:CORPTSMA	2656	L18874	Brevibacterium lactofermentum phosphoenolpyruvate sugar phosphotransferase (ptsG) gene, complete cds.		Corynebacterium glutamicum	100,000	24-Nov-94
rx01944	1095	GB_HTG2:AC004840	162485	AC004840	complete cds.		Corynebacterium glutamicum	37,008	24-Nov-94
		GB_HTG2:AC004840	162485	AC004840	Corynebacterium glutamicum phosphoenolpyruvate sugar phosphotransferase (ptsM) mRNA, complete cds.		Homo sapiens	36,287	12-Jun-98
		GB_V1:MCU60315	190289	U60315	Homo sapiens clone DJ0607J02, *** SEQUENCING IN PROGRESS ***	12	Homo sapiens	36,287	12-Jun-98
rx01945	2115	GB_GSS15:AQ632158	445	AQ632158	Homo sapiens clone DJ0607J02, *** SEQUENCING IN PROGRESS ***	12	Molluscum contagiosum	37,650	17-Aug-96
		GB_GSS15:AQ632158	445	AQ632158	Molluscum contagiosum virus subtype 1, complete genome.		virus subtype 1	38,085	17-Jun-99
		GB_EST16:C35275	300	C35275	RPCI-11-473F7, TJ RPCI-11 Homo sapiens genomic clone RPCI-11-473F7, genomic survey sequence.		Homo sapiens	40,275	17-Jun-99
rx01957	585	GB_IN2:CELC29H12	42635	U23169	RPCI-11-473F7, TJ RPCI-11 Homo sapiens genomic clone RPCI-11-473F7, genomic survey sequence.		Caenorhabditis elegans	37,698	18-OCT-1999
		GB_GSS1:CNS00GP0	1101	AL072364	C35275 Yuji Kohara unpublished cDNA: Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk427e6 3', mRNA sequence.		Caenorhabditis elegans	34,477	13-Jul-95
rx01958	630	GB_PR1:HSIL1RECA	12565	X64532	Caenorhabditis elegans cosmid C29H12.		Drosophila melanogaster	34,321	3-Jun-99
		GB_PR3:HSU65590	33414	U65590	Drosophila melanogaster genome survey sequence T7 end of BAC: BACR33108 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		Homo sapiens	42,079	25-Jun-97
		GB_GSS11:AQ293877	476	AQ293877	H sapiens gene for interleukin-1 receptor antagonist.		Homo sapiens	42,079	21-DEC-1997
rx01959					Homo sapiens IL-1 receptor antagonist IL-1Ra (IL-1RN) gene, alternatively spliced forms, complete cds.		Homo sapiens	36,772	15-DEC-1998
rx01960					HS_2254_A2_C03_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2254 Col=6 Row=E, genomic survey sequence.				

Attorney Work Product  
Not for Distribution  
This document contains information that is confidential and may be subject to the attorney-client privilege. It is to be used only for the purposes for which it was prepared and is not to be disseminated to any other person without the prior written consent of the undersigned.

**TABLE 4: ALIGNMENT RESULTS**

rxa01961	603	GB_BA2:U67460	12589	U67460	Methanococcus jannaschii section 2 of 150 of the complete genome.	Methanococcus jannaschii	36,013	28-Jan-98
		GB_EST25:A1260761	626	A1260761	LP04729 3prime LP Drosophila melanogaster larval-early pupal pOT2 Drosophila melanogaster cDNA clone LP04729 3prime, mRNA sequence.	Drosophila melanogaster	39,130	17-Nov-98
		GB_EST25:A1294890	554	A1294890	LP08371 3prime LP Drosophila melanogaster larval-early pupal pOT2 Drosophila melanogaster cDNA clone LP08371 3prime, mRNA sequence.	Drosophila melanogaster	38,940	01-DEC-1998
rxa01962	693	GB_BA2:U67460	12589	U67460	Methanococcus jannaschii section 2 of 150 of the complete genome.	Methanococcus jannaschii	34,795	28-Jan-98
		GB_BA2:U67460	12589	U67460	Methanococcus jannaschii section 2 of 150 of the complete genome.	Methanococcus jannaschii	37,666	28-Jan-98
rxa01963	784	GB_PR4:AC006397	91460	AC006397	Homo sapiens BAC clone GS170102 from 7p21-p15.1, complete sequence.	Homo sapiens	35,401	17-Feb-99
		GB_PR3:AC005565	39441	AC005565	Homo sapiens chromosome 16, cosmid clone 444B9 (LANL), complete sequence.	Homo sapiens	37,891	30-Aug-98
		GB_PR2:D86999	40778	D86999	Homo sapiens immunoglobulin lambda gene locus DNA, clone:22A12.	Homo sapiens	38,212	5-Nov-99
rxa01964	1426	GB_GSS13:AQ465174	437	AQ465174	HS_5109_A2_C07_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=685 Col=14 Row=E, genomic survey sequence.	Homo sapiens	40,084	23-Apr-99
		GB_GSS13:AQ457528	599	AQ457528	HS_5087_A1_H06_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=663 Col=11 Row=O, genomic survey sequence.	Homo sapiens	36,728	23-Apr-99
		GB_GSS13:AQ454028	527	AQ454028	HS_5171_B1_D12_TTA RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=747 Col=23 Row=H, genomic survey sequence.	Homo sapiens	37,137	21-Apr-99
rxa01965	684	GB_IN2:AC004283	80095	AC004283	Drosophila melanogaster DNA sequence (P1 DS05557 (D152)), complete sequence.	Drosophila melanogaster	36,982	29-Aug-98
		GB_EST28:A1541584	585	A1541584	SD02734 5prime SD Drosophila melanogaster Schneider L2 cell culture pOT2 Drosophila melanogaster cDNA clone SD02734 5prime, mRNA sequence.	Drosophila melanogaster	38,375	22-MAR-1999
		GB_EST22:A1013385	434	A1013385	EST208060 Normalized rat spleen, Bentio Soares Rattus sp. cDNA clone RSPBK32 3' end, mRNA sequence.	Rattus sp.	42,126	31-Jan-99
rxa01966	780	GB_GSS1:CNS00AMI	810	AL055260	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR21F10 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	Drosophila melanogaster	37,592	4-Jun-99
		GB_HTG3:AC009567	164147	AC009567	Homo sapiens chromosome 4 clone 21_G_20 map 4, *** SEQUENCING IN PROGRESS *** 14 unordered pieces.	Homo sapiens	35,844	29-Sep-99
		GB_HTG3:AC009567	164147	AC009567	Homo sapiens chromosome 4 clone 21_G_20 map 4, *** SEQUENCING IN PROGRESS *** 14 unordered pieces.	Homo sapiens	35,844	29-Sep-99
rxa01968	1281	GB_PR4:AC005146	157653	AC005146	Homo sapiens 12q24.2 PAC RPC11-157K6 (Roswell Park Cancer Institute Human PAC library) complete sequence.	Homo sapiens	40,620	11-Nov-98
		GB_PR4:AC006549	174844	AC006549	Homo sapiens, complete sequence.	Homo sapiens	34,202	26-Nov-99
rxa01969	489	GB_PR2:HSAC002070	165197	AC002070	Human BAC clone 7E17 from 12q, complete sequence.	Homo sapiens	36,482	12-MAY-1997
		GB_SY:SCUS3587	4546	U53587	Artificial Corynebacterium glutamicum IS1207-derived transposon transposase genes, complete cds, and 3'-aminoglycoside phosphotransferase (aphA-3) gene, complete cds.	synthetic construct	98,160	06-MAY-1996
		GB_PAT:143826	1452	I43826	Sequence 1 from patent US 5633154.	Unknown.	98,963	07-OCT-1997
rxa01973		GB_BA1:CGIS1207	1453	X96962	C. glutamicum insertion sequence IS1207 and transposase gene.	Corynebacterium glutamicum	98,755	01-MAY-1997
rxa01974	1908	GB_PL1:SC9320A	24000	Z68329	S.cerevisiae chromosome IV cosmid 9320A.	Saccharomyces cerevisiae	36,528	11-Aug-97
		GB_PL1:SC9320X	22253	Z70202	S.cerevisiae chromosome IV cosmid 9320X.	Saccharomyces cerevisiae	36,528	11-Aug-97
		GB_PR4:AC006210	186986	AC006210	Homo sapiens Xp22-150 BAC GSHB-309P15 (Genome Systems Human BAC Library) complete sequence.	Homo sapiens	38,351	31-DEC-1998

TABLE 4: ALIGNMENT RESULTS

rx01976	1644	GB_BA2:CGU13922	4412	U13922	Corynebacterium glutamicum putative type II 5-cytosine methyltransferase (cgIM) and putative type II restriction endonuclease (cgIR) and putative type I or type III restriction endonuclease (cgIR) genes, complete cds.	Corynebacterium glutamicum	47,727	3-Feb-98
		GB_BA2:CGU13922	4412	U13922	Corynebacterium glutamicum putative type II 5-cytosine methyltransferase (cgIM) and putative type II restriction endonuclease (cgIR) and putative type I or type III restriction endonuclease (cgIR) genes, complete cds.	Corynebacterium glutamicum	37,515	3-Feb-98
rx01977	699	GB_PR3:HS426F10	91840	AL023586	Human DNA sequence from clone 426F10 on chromosome 1p36.21-36.33 Contains EST, CA repeat, STS, complete sequence.	Homo sapiens	38,295	23-Nov-99
		GB_GSS6:AQ823696	587	AQ823696	HS_3228_A1_D10_TTC CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3228 Col=19 Row=G, genomic survey sequence.	Homo sapiens	37,637	26-Aug-99
		GB_VI:HS3DNA1	892	M29629	Varicella-zoster virus (VZV) DNA, 5' end.	Human herpesvirus 3	33,857	2-Aug-93
rx01982	1281	GB_EST8:AA009257	472	AA009257	mh02g07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:441372 5', mRNA sequence.	Mus musculus	36,017	26-Jul-96
		GB_EST8:AA009257	472	AA009257	mh02g07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:441372 5', mRNA sequence.	Mus musculus	38,710	26-Jul-96
rx01987	336	GB_PR4:AC009330	174768	AC009330	Homo sapiens clone RP11-83M17 from 7q31, complete sequence.	Homo sapiens	33,036	09-DEC-1999
		GB_HTG6:AC007300	171472	AC007300	Drosophila melanogaster chromosome 2 clone BACR0915 (D570) RPC1-98 09.115 map 32A-32A strain Y, cn bw sp. *** SEQUENCING IN PROGRESS ***. 43 unordered pieces.	Drosophila melanogaster	38,782	07-DEC-1999
rx01988	549	GB_HTG3:AC009542	160367	AC009542	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS ***. 17 unordered pieces.	Homo sapiens	35,065	27-Aug-99
		GB_IN1:AC002446	50978	AC002446	Drosophila melanogaster (P1 D506754 (D83)) DNA sequence, complete sequence.	Drosophila melanogaster	38,920	19-Aug-97
		GB_PR3:HSU19H10	43303	AL021182	Human DNA sequence from cosmid U19H10 on chromosome X. Contains ESTs and CA repeat.	Homo sapiens	39,006	23-Nov-99
		GB_PR3:HSU19H10	43303	AL021182	Human DNA sequence from cosmid U19H10 on chromosome X. Contains ESTs and CA repeat.	Homo sapiens	38,447	23-Nov-99
rx01990	1022	GB_GSS9:AQ100628	382	AQ100628	HS_3054_B2_A05_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3054 Col=10 Row=B, genomic survey sequence.	Homo sapiens	38,320	27-Aug-98
		GB_EST14:AA0406679	338	AA0406679	MBAFCZ2F12T3A Brugia malayi adult female cDNA (SAW96MLW-BnAF) Brugia malayi cDNA clone AFCZ2F12 5', mRNA sequence.	Brugia malayi	38,039	01-MAY-1997
		GB_EST21:AA991134	600	AA991134	BSBmMFSZ06C1SK Brugia malayi microfilaria cDNA (SAW94LS-BmM) Brugia malayi cDNA clone AFCZ2F12 5', mRNA sequence.	Brugia malayi	40,090	5-Jun-98
rx01991	996	GB_HTG3:AC007559	156374	AC007559	Homo sapiens clone NH0364J06, *** SEQUENCING IN PROGRESS ***. 29 unordered pieces.	Homo sapiens	35,931	13-Aug-99
		GB_HTG3:AC007559	156374	AC007559	Homo sapiens clone NH0364J06, *** SEQUENCING IN PROGRESS ***. 29 unordered pieces.	Homo sapiens	35,931	13-Aug-99
		GB_HTG3:AC007559	156374	AC007559	Homo sapiens clone NH0364J06, *** SEQUENCING IN PROGRESS ***. 29 unordered pieces.	Homo sapiens	36,120	13-Aug-99
rx01992								
rx01996	660	GB_BA1:XCXPSGEN	5324	X59079	X.campestris xps E, F, G, H, I, and J genes for protein secretion and pathogenicity functions. Sheared DNA-5C7.TR Sheared DNA Trypanosoma brucei genomic clone Sheared DNA-5C7, genomic survey sequence.	Xanthomonas campestris Trypanosoma brucei	38,447 49,180	12-Sep-93 23-Jun-99
		GB_GSS4:AQ741049	713	AQ741049	HS_2272_A2_B02_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2272 Col=4 Row=C, genomic survey sequence.	Homo sapiens	39,614	16-Jul-99
rx01999	594	GB_HTG2:AC007321	159558	AC007321	Homo sapiens clone NH0507C01, *** SEQUENCING IN PROGRESS ***. 4 unordered pieces.	Homo sapiens	34,524	16-Apr-99
		GB_HTG2:AC007321	159558	AC007321	Homo sapiens clone NH0507C01, *** SEQUENCING IN PROGRESS ***. 4 unordered pieces.	Homo sapiens	34,524	16-Apr-99
		GB_HTG2:AC007321	159558	AC007321	Homo sapiens clone NH0507C01, *** SEQUENCING IN PROGRESS ***. 4 unordered pieces.	Homo sapiens	35,094	16-Apr-99
rx02001	1281	GB_BA1:D90917	154619	D90917	Synechocystis sp. PCC6803 complete genome, 27/27, 3418852-3573470.	Synechocystis sp.	50,474	7-Feb-99
		GB_BA1:MTV016	53662	AL021841	Mycobacterium tuberculosis H37Rv complete genome, segment 143/162.	Mycobacterium tuberculosis	39,223	23-Jun-99
rx02004	230	GB_BA1:BSUB0006	210440	Z99109	Bacillus subtilis complete genome (section 6 of 21): from 999501 to 1209940.	Bacillus subtilis	36,445	26-Nov-97
		GB_HTG3:AC011491	162134	AC011491	Homo sapiens chromosome 19 clone C17978SKB_180A7, *** SEQUENCING IN PROGRESS ***. 161 unordered pieces.	Homo sapiens	37,383	07-OCT-1999



TABLE 4: ALIGNMENT RESULTS

GB_HTG3:AC011491	162134	AC011491	Homo sapiens chromosome 19 clone CIT978SKB_180A7, *** SEQUENCING IN PROGRESS ***, 161 unordered pieces.	Homo sapiens	37,383	07-OCT-1999
GB_HTG3:AC011357	160876	AC011357	Homo sapiens chromosome 5 clone CIT-HSPC_362D12, *** SEQUENCING IN PROGRESS ***, 42 unordered pieces.	Homo sapiens	54,585	06-OCT-1999
GB_HTG2:AC006901	294136	AC006901	Caenorhabditis elegans clone Y74A11X, *** SEQUENCING IN PROGRESS ***, 81 unordered pieces.	Caenorhabditis elegans	37,500	26-Feb-99
GB_HTG2:AC006901	294136	AC006901	Caenorhabditis elegans clone Y74A11X, *** SEQUENCING IN PROGRESS ***, 81 unordered pieces.	Caenorhabditis elegans	37,500	26-Feb-99
GB_HTG1:CEY70C5	224525	Z98878	Caenorhabditis elegans chromosome V clone Y70C5, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Caenorhabditis elegans	36,021	28-Aug-97
GB_BA2:CGLO12293	2952	AJ012293	Corynebacterium glutamicum ilvD gene.	Corynebacterium glutamicum	100,000	01-OCT-1999
GB_PR4:AC006285	150172	AC006285	Homo sapiens, complete sequence.	Homo sapiens	38,740	15-Nov-99
GB_PR4:AC006285	150172	AC006285	Homo sapiens, complete sequence.	Homo sapiens	38,701	15-Nov-99
GB_PR4:AC004926	153556	AC004926	Homo sapiens PAC clone DJ0910H09 from 7q21.1-q21.2, complete sequence.	Homo sapiens	38,462	17-Feb-99
GB_PR2:AP000053	100000	AP000053	Homo sapiens genomic DNA, chromosome 21q22.1, segment 24/28, complete sequence.	Homo sapiens	37,674	20-Nov-99
GB_PR2:AP000121	100000	AP000121	Homo sapiens genomic DNA of 21q22.1, GART and AML related, SLC5A3-14A4 region, segment 4/8, complete sequence.	Homo sapiens	37,674	25-Sep-99
GB_BA1:MTCI237	27030	Z94752	Mycobacterium tuberculosis H37Rv complete genome, segment 46/162.	Mycobacterium tuberculosis	39,308	17-Jun-98
GB_HTG4:AC010885	201581	AC010885	Homo sapiens chromosome unknown clone NH0368K23, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	37,206	29-OCT-1999
GB_HTG4:AC010885	201581	AC010885	Homo sapiens chromosome unknown clone NH0368K23, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	37,206	29-OCT-1999
GB_HTG3:AC010769	119431	AC010769	Homo sapiens chromosome 15 clone 28_B_17 map 15, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	33,816	22-Sep-99
GB_HTG3:AC010769	119431	AC010769	Homo sapiens chromosome 15 clone 28_B_17 map 15, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	33,816	22-Sep-99
GB_EST34:A1783738	320	A1783738	tu45b07 x1 NCI_CGAP_P128 Homo sapiens cDNA clone IMAGE:2253973 3', mRNA sequence.	Homo sapiens	37,855	1-Jul-99
GB_EST15:AA473289	318	AA473289	vd44g09.1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:803488 5', mRNA sequence.	Mus musculus	38,153	18-Jun-97
GB_HTG1:CEY102A5_3110000	Z99711	Z99711	Caenorhabditis elegans chromosome V clone Y102A5, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Caenorhabditis elegans	38,867	9-Jun-98
GB_HTG1:CEY102A5_3110000	Z99711	Z99711	Caenorhabditis elegans chromosome V clone Y102A5, *** SEQUENCING IN PROGRESS ***, in unordered pieces.	Caenorhabditis elegans	38,867	9-Jun-98
GB_BA1:CGDNAAROP 2612	X85965	X85965	C. glutamicum ORF3 and aroP gene.	Corynebacterium glutamicum	99,905	30-Nov-97
GB_BA1:CGDAPE 1966	X81379	X81379	C. glutamicum dapE gene and orf2.	Corynebacterium glutamicum	36,406	8-Aug-95
GB_BA1:SCI7 34893	AL096743	AL096743	Streptomyces coelicolor cosmid 17.	Streptomyces coelicolor	36,782	1-Jul-99
GB_BA1:CGDAPE 1966	X81379	X81379	C. glutamicum dapE gene and orf2.	Corynebacterium glutamicum	99,864	8-Aug-95
GB_PR3:AC004067	161326	AC004067	Homo sapiens chromosome 4 clone B366O24 map 4q25, complete sequence.	Homo sapiens	39,651	8-Nov-98
GB_HTG3:AC009725	145005	AC009725	Mus musculus chromosome 6 clone 388_N_17 map 6, *** SEQUENCING IN PROGRESS ***, 2 ordered pieces.	Mus musculus	38,215	01-OCT-1999
GB_PR4:AC006961	171419	AC006961	Homo sapiens chromosome 18, clone RP11-31P16, complete sequence.	Homo sapiens	37,703	19-Nov-99
GB_PR4:AC006961	171419	AC006961	Homo sapiens chromosome 18, clone RP11-31P16, complete sequence.	Homo sapiens	40,839	19-Nov-99
GB_PR2:HSDJ799G3	127639	AL078624	Human DNA sequence from clone RPA-799G3 on chromosome 1q42.11-42.3, complete sequence.	Homo sapiens	35,657	22-Nov-99
GB_BA1:AOPCZA361	37941	AJ223998	Amycolatopsis orientalis cosmid PCZA361.	Amycolatopsis orientalis	46,341	29-MAR-1999
GB_PL1:ATDNADAL1 8747	Y14851	Y14851	Arabidopsis thaliana dal1 gene.	Arabidopsis thaliana	35,223	23-Sep-97
GB_PR3:AC005609	157970	AC005609	Homo sapiens chromosome 5, BAC clone 203o13 (LBNL H155), complete sequence.	Homo sapiens	37,176	4-Sep-98

**TABLE 4: ALIGNMENT RESULTS**

rx02040	551	GB_HTG3:AC008468	245016	AC008468	Homo sapiens chromosome 5 clone CIT-HSPC_365B8, *** SEQUENCING IN PROGRESS ***; 14 unordered pieces.	Homo sapiens	36,842	3-Aug-99
		GB_HTG3:AC008468	245016	AC008468	Homo sapiens chromosome 5 clone CIT-HSPC_365B8, *** SEQUENCING IN PROGRESS ***; 14 unordered pieces.	Homo sapiens	36,842	3-Aug-99
		GB_HTG3:AC009303	198549	AC009303	Homo sapiens clone NH0098C01, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	35,424	13-Aug-99
		GB_HTG3:AC009303	198549	AC009303	Homo sapiens clone NH0098C01, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	35,424	13-Aug-99
		GB_HTG3:AC009303	198549	AC009303	Homo sapiens clone NH0098C01, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	32,597	13-Aug-99
rx02045	384	GB_EST3:R23812	459	R23812	YH34G05.11 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:131672 5', mRNA sequence.	Homo sapiens	47,273	20-Apr-95
		GB_EST19:AA778733	611	AA778733	af8c02.s1 Soares testis_NHT Homo sapiens cDNA clone 1049091 3', mRNA sequence.	Homo sapiens	38,806	5-Feb-98
		GB_PR2:HSU59185	2529	U59185	Human putative monocarboxylate transporter (MCT) mRNA, complete cds.	Homo sapiens	43,164	03-OCT-1997
rx02046	540	GB_HTG3:AC005507	196595	AC005507	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Plasmodium falciparum	37,218	23-Sep-99
		GB_HTG3:AC005507	196595	AC005507	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Plasmodium falciparum	37,218	23-Sep-99
		GB_EST1:TO9984	330	TO9984	0540m7 gmbPHB3.1, G. Roman Reddy Plasmodium falciparum genomic clone 0540m, mRNA sequence.	Plasmodium falciparum	32,121	29-Nov-93
rx02049	732	GB_PL1:ATHCOLR	1088	M37778	A.thaliana middle repetitive sequence.	Arabidopsis thaliana	33,755	27-Apr-93
		GB_PL2:ATAC006954	87035	AC006954	Arabidopsis thaliana chromosome II BAC F25P17 genomic sequence, complete sequence.	Arabidopsis thaliana	33,755	7-Apr-99
		GB_EST33:A1779087	495	A1779087	EST259946 tomato susceptible, Cornell Lycopersicon esculentum cDNA clone cLES7C7, mRNA sequence.	Lycopersicon esculentum	35,223	29-Jun-99
rx02050	918	GB_HTG2:AC008307	91654	AC008307	Drosophila melanogaster chromosome 3 clone BACR03D22 (D709) RPCI-98 03.D.22 map 86F-87A strain y, cn bw sp, *** SEQUENCING IN PROGRESS ***; 94 unordered pieces.	Drosophila melanogaster	36,182	2-Aug-99
		GB_HTG2:AC008307	91654	AC008307	Drosophila melanogaster chromosome 3 clone BACR03D22 (D709) RPCI-98 03.D.22 map 86F-87A strain y, cn bw sp, *** SEQUENCING IN PROGRESS ***; 94 unordered pieces.	Drosophila melanogaster	36,182	2-Aug-99
		GB_EST31:AU060923	663	AU060923	AU060923 Dictyostelium discoideum SL (H.Urushiara) Dictyostelium discoideum cDNA clone SLC248, mRNA sequence.	Dictyostelium discoideum	39,388	20-MAY-1999
rx02051	621	GB_EST16:C43896	369	C43896	C43896 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk336c10 5', mRNA sequence.	Caenorhabditis elegans	41,160	18-OCT-1999
		GB_EST16:C40413	360	C40413	C40413 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk230e7 5', mRNA sequence.	Caenorhabditis elegans	42,090	18-OCT-1999
		GB_EST36:AV191515	360	AV191515	AV191515 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk594g6 5', mRNA sequence.	Caenorhabditis elegans	40,833	22-Jul-99
rx02053	702	GB_INT:AF035284	6397	AF035284	Drosophila melanogaster POU domain protein (pdm-1) gene, promoter region and exon 1.	Drosophila melanogaster	36,888	04-DEC-1997
		GB_BA1:MTV036	24055	AL021931	Mycobacterium tuberculosis H37Rv complete genome, segment 19/162.	Mycobacterium tuberculosis	37,647	17-Jun-98
		GB_EST30:AI658116	497	AI658116	fc22e07.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA 5' similar to TR:Q15883 Q15883 X104.1, mRNA sequence.	Danio rerio	41,129	06-MAY-1999
rx02057	654	GB_PAT:E14601	4394	E14601	Brevibacterium lactofermentum gene for alpha-ketoglutaric acid dehydrogenase.	Corynebacterium glutamicum	96,667	28-Jul-99
		GB_BA1:D84102	4394	D84102	Corynebacterium glutamicum DNA for 2-oxoglutarate dehydrogenase, complete cds.	Corynebacterium glutamicum	96,667	6-Feb-99
		GB_PAT:E14601	4394	E14601	Brevibacterium lactofermentum gene for alpha-ketoglutaric acid dehydrogenase.	Corynebacterium glutamicum	38,199	28-Jul-99
rx02058	675	GB_GSS9:AQ129748	432	AQ129748	HS_2254_B1_H03_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2254 Col=5 Row=P, genomic survey sequence.	Homo sapiens	40,000	23-Sep-98
		GB_PL2:T15B16	90596	AF104919	Arabidopsis thaliana BAC T15B16.	Arabidopsis thaliana	35,338	11-Nov-98
		GB_PL2:AC007138	120185	AC007138	Arabidopsis thaliana BAC TT811 from chromosome IV near 10 cM, complete sequence.	Arabidopsis thaliana	39,426	1-Apr-99
rx02059	618	GB_HTG3:AC008672	131573	AC008672	Homo sapiens chromosome 5 clone CIT978SKB_3B12, *** SEQUENCING IN PROGRESS ***; 71 unordered pieces.	Homo sapiens	40,871	3-Aug-99

**TABLE 4: ALIGNMENT RESULTS**

GB_HTC3:AC008672	131573	AC008672	Homo sapiens chromosome 5 clone C1978SKB_3B12, *** SEQUENCING IN PROGRESS ***	71	Homo sapiens	40,871	3-Aug-99
GB_IN1:DDU06228	2695	U06228	Dictyostellium discoideum CRAC (dagA) gene, complete cds.		Dictyostellium discoideum	36,513	1-Feb-95
GB_PR2:CNS01DRA	198444	AL110505	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-816J8 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.		Homo sapiens	36,903	11-Nov-99
GB_PR2:HS230G1	125515	Z84466	Homo sapiens DNA sequence from PAC 230G1 on chromosome Xp11.3. Contains EST, STS and GSS, complete sequence.		Homo sapiens	41,751	23-Nov-99
GB_PR2:CNS01DRA	198444	AL110505	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-816J8 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.		Homo sapiens	36,569	11-Nov-99
GB_HTC3:AC009543	159209	AC009543	Homo sapiens chromosome 11 clone 63_H_13 map 11, *** SEQUENCING IN PROGRESS ***	12	Homo sapiens	35,366	07-OCT-1999
GB_HTC3:AC009543	159209	AC009543	unordered pieces.		Homo sapiens	35,366	07-OCT-1999
GB_PR2:HS230G1	125515	Z84466	Homo sapiens DNA sequence from PAC 230G1 on chromosome Xp11.3. Contains EST, STS and GSS, complete sequence.		Homo sapiens	36,348	23-Nov-99
GB_BA1:MTC161	13540	Z98260	Mycobacterium tuberculosis H37Rv complete genome, segment 53/162.		Mycobacterium tuberculosis	36,642	17-Jun-98
GB_BA1:MTC161	13540	Z98260	Mycobacterium tuberculosis H37Rv complete genome, segment 53/162.		Mycobacterium tuberculosis	39,474	17-Jun-98
GB_BA2:AE001112	14704	AE001112	Archaeoglobus fulgidus section 167 of 172 of the complete genome.		Archaeoglobus fulgidus	39,352	15-DEC-1997
GB_PR3:HSDJ47M23	129320	AL096816	Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone dJ47M23, complete sequence.		Homo sapiens	40,413	23-Nov-99
GB_HTC3:AC009139	152666	AC009139	Homo sapiens chromosome 16 clone RPCI-11_538112, *** SEQUENCING IN PROGRESS ***	27	Homo sapiens	35,275	3-Aug-99
GB_EST34:A1789138	594	A1789138	unordered pieces.		Mus musculus	43,041	2-Jul-99
GB_GSS1:FR0013995	552	AL005239	uk51e02.y1 Sugano mouse kidney mRna Mus musculus cDNA clone IMAGE:1972538 5' similar to WP:R10H10.7 CE08298 TYROSINE-PROTEIN KINASE LIKE ; mRNA sequence.		Fugu rubripes	37,079	18-Sep-97
GB_GSS3:B27548	642	B27548	F. rubripes GSS sequence, clone 137018aC6, genomic survey sequence.		Arabidopsis thaliana	37,056	13-OCT-1997
GB_GSS15:AQ652136	411	AQ652136	F19J9TFB IGF Arabidopsis thaliana genomic clone F19J9, genomic survey sequence.		Trypanosoma brucei	39,259	22-Jun-99
GB_GSS15:AQ652498	450	AQ652498	Sheared DNA-7M23.TR Sheared DNA Trypanosoma brucei genomic clone Sheared DNA-7M23, genomic survey sequence.		Trypanosoma brucei	39,597	22-Jun-99
GB_HTC3:AC009683	171597	AC009683	Sheared DNA-22K16.TR Sheared DNA Trypanosoma brucei genomic clone Sheared DNA-22K16, genomic survey sequence.		Homo sapiens	37,053	29-Sep-99
GB_OV:BSU43200	2070	U43200	Homo sapiens chromosome 8 clone 76_N_5 map 8, *** SEQUENCING IN PROGRESS ***	10	Boreogadus saida	37,016	12-MAY-1997
GB_BA2:AE001615	12401	AE001615	unordered pieces.		Chlamydia pneumoniae	36,385	08-MAR-1999
GB_OV:BSU43200	2070	U43200	Boreogadus saida antifreeze glycoprotein precursor gene, complete cds.		Boreogadus saida	37,672	12-MAY-1997
GB_PR4:AC007382	80547	AC007382	Homo sapiens clone NH0288C18, complete sequence.		Chlamydia pneumoniae	41,138	8-Sep-99
GB_GSS4:AQ731530	508	AQ731530	HS_5543_A2_A05_TTA RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1119 Col=10 Row=A, genomic survey sequence.		Homo sapiens	36,245	15-Jul-99
GB_PR4:AC007382	80547	AC007382	Homo sapiens clone NH0288C18, complete sequence.		Homo sapiens	35,398	8-Sep-99
GB_PR2:AP000031	149298	AP000031	Homo sapiens genomic DNA, chromosome 21q22.1, segment 2/28, complete sequence.		Homo sapiens	38,699	20-Nov-99
GB_PR2:AP000135	111894	AP000135	Homo sapiens genomic DNA of 21q22.1, GART and AML, 143D11-11988 region, segment 10/10, complete sequence.		Homo sapiens	38,699	25-Sep-99
GB_PR2:AP000213	100000	AP000213	Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone f43D11-11988, segment 11/12, complete sequence.		Homo sapiens	38,699	20-Nov-99
GB_PL2:ATAC006841	123183	AC006841	Arabidopsis thaliana chromosome II BAC F3K23 genomic sequence, complete sequence.		Arabidopsis thaliana	39,826	6-Apr-99
GB_PL1:SCXV55KB	54719	Z70678	S. cerevisiae chromosome XV DNA, 54.7 kb region.		Saccharomyces cerevisiae	37,817	16-MAY-1997

TABLE 4: ALIGNMENT RESULTS

rx02091	774	GB_PR1:D87675	301692	D87675	Homo sapiens DNA for amyloid precursor protein, complete cds.	Homo sapiens	39,140	22-Sep-97
		GB_BA2:AF031929	2675	AF031929	Lactobacillus helveticus co-chaperonin GroES and chaperonin GroEL genes, complete cds; and DNA mismatch repair enzyme (hexA) gene, partial cds.	Lactobacillus helveticus	35,509	8-Aug-98
		GB_BA1:CGU43536	3464	U43536	Corynebacterium glutamicum heat shock, ATP-binding protein (cibp) gene, complete cds.	Corynebacterium glutamicum	39,124	13-MAR-1997
rx02094	840	GB_BA1:RCFBC	3874	X03476	Rhodopseudomonas sphaeroides fbc operon (fbcF, fbcB, fbcC genes).	Rhodobacter sphaeroides	36,478	10-Feb-99
		GB_RO:MMHC188A7	120990	AF109719	Mus musculus casein kinase 2 beta subunit (gMCK2) gene, partial cds; BAT4, NG20 (NG20), BAT3, BAT2, AIF-1, B144, lymphotoxin beta, TNF, and TNF beta genes, complete cds; IKBL gene, partial cds; and unknown gene.	Mus musculus	38,118	25-OCT-1999
		GB_EST10:AA168948	549	AA168948	ms39c08.r1 Life Tech mouse embryo 13 5dpc 10666014 Mus musculus cDNA clone IMAGE:613934 5' similar to gb:U15980 Mus musculus (MOUSE); mRNA sequence.	Mus musculus	40,789	19-DEC-1996
		GB_EST24:AI101223	348	AI101223	EST210512 Normalized rat brain, Bento Soares Rattus sp. cDNA clone RBRBK04 3' end, mRNA sequence.	Rattus sp.	40,634	31-Jan-99
rx02097	3495	GB_BA2:RMU31512	1974	U31512	Rhizobium meliloti RmDEGP (degP) gene, complete cds.	Sinorhizobium meliloti	39,159	05-MAR-1996
		GB_EST15:AA502050	346	AA502050	ng57c04.s1 NCI_CGAP_Lip2 Homo sapiens cDNA clone IMAGE:938886 similar to contains Alu repetitive element; mRNA sequence.	Homo sapiens	40,751	18-Aug-97
		GB_EST16:AA589883	400	AA589883	vi80h08.r1 Stratagene mouse diaphragm (#937303) Mus musculus cDNA clone IMAGE:987039 5' similar to gb:M75716 Mus musculus alpha-1 protease inhibitor 2 (MOUSE); mRNA sequence.	Mus musculus	35,427	16-Sep-97
rx02102	1281	GB_IN2:AF135118	10830	AF135118	Drosophila melanogaster laminin alpha1.2 (wing blister) mRNA, complete cds.	Drosophila melanogaster	36,926	23-MAY-1999
		GB_IN1:AC002516	48158	AC002516	Drosophila melanogaster (P1 DS01068 (D37)) DNA sequence, complete sequence.	Drosophila melanogaster	37,838	28-Aug-97
		GB_GSS3:B09866	1194	B09866	T16G20-Sp6 TAMU Arabidopsis thaliana genomic clone T16G20, genomic survey sequence.	Arabidopsis thaliana	35,332	14-MAY-1997
rx02103	519	GB_GSS3:B09866	1194	B09866	T16G20-Sp6 TAMU Arabidopsis thaliana genomic clone T16G20, genomic survey sequence.	Arabidopsis thaliana	36,310	14-MAY-1997
		GB_EST34:AV165661	290	AV165661	AV165661 Mus musculus head C57BL/6J 13-day embryo Mus musculus cDNA clone 3110038F04, mRNA sequence.	Mus musculus	37,716	6-Jul-99
rx02104	1245	GB_RO:AC002121	84056	AC002121	Genomic sequence from Mouse 11, complete sequence.	Mus musculus	39,096	10-Jul-97
		GB_BA2:ECOUW67_0	110000	U18997	Escherichia coli K-12 chromosomal region from 67.4 to 76.0 minutes.	Escherichia coli	39,024	22-Dec-94
		GB_BA2:AE000394	12221	AE000394	Escherichia coli K-12 MG1655 section 284 of 400 of the complete genome.	Escherichia coli	39,024	12-Nov-98
rx02107		GB_BA1:ECORNPBW	4434	D90212	E. coli mpB gene and ORFs.	Escherichia coli	45,255	7-Feb-99
rx02108	732	GB_BA1:D90912	128598	D90912	Synechocystis sp. PCC6803 complete genome, 14/27, 1719644-1848241.	Synechocystis sp.	38,095	7-Feb-99
		GB_GSS13:AQ498890	431	AQ498890	HS_5154_A1_E06_SPEE RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=730 Col=11 Row=I, genomic survey sequence.	Homo sapiens	37,150	28-Apr-99
rx02109	1044	GB_BA1:D90912	128598	D90912	Synechocystis sp. PCC6803 complete genome, 14/27, 1719644-1848241.	Synechocystis sp.	38,781	7-Feb-99
		GB_IN2:AC001658	91019	AC001658	Drosophila melanogaster DNA sequence (P1 DS00913 (D24)), complete sequence.	Drosophila melanogaster	35,687	17-Jul-98
		GB_EST19:AA802304	581	AA802304	GM04170 5prime GM Drosophila melanogaster ovary BlueScript Drosophila melanogaster cDNA clone GM04170 5prime, mRNA sequence.	Drosophila melanogaster	40,451	25-Nov-98
		GB_EST35:A1822653	733	A1822653	LO-1173T3 Ice plant Lambda Uni-Zap XR expression library, 0 hours NaCl treatment Mesembryanthemum crystallinum cDNA clone LO-1173 5' similar to Profilin 1 (AF092547) [Ricinus communis, mRNA sequence.	Mesembryanthemum crystallinum	36,957	12-Jul-99
rx02114	509	GB_BA1:SMU94899	2379	U94899	Sinorhizobium meliloti dissimilatory nitrous oxide reduction proteins NosY, NosL and NosX genes, complete cds.	Sinorhizobium meliloti	42,744	6-Sep-97
		GB_BA1:SMU94899	2379	U94899	Sinorhizobium meliloti dissimilatory nitrous oxide reduction proteins NosY, NosL and NosX genes, complete cds.	Sinorhizobium meliloti	39,486	6-Sep-97
rx02117	597	GB_BA2:AF094575	18754	AF094575	Streptococcus pneumoniae serotype 19A, DexB (dexB) gene, partial sequence; capsular polysaccharide biosynthesis operon, complete sequence; and oligopeptide permease AIIA (aiiA) gene, partial cds.	Streptococcus pneumoniae	40,480	9-Sep-99

TABLE 4: ALIGNMENT RESULTS

GB_BA2:AF105113	5832	AF105113	Streptococcus pneumoniae type 19A putative oligosaccharide repeat unit transporter (cps19AJ) gene, partial cds; UDP-N-acetyl glucosamine-2-epimerase (cps19AK), glucose-1-phosphate thymidyl transferase (cps19AL), dTDP-4-keto-6-deoxyglucose-3,5-epimerase (cps19AM), dTDP-glucose-4,6 dehydratase (cps19AN), and dTDP-L-rhamnose synthase (cps19AO) genes, complete cds; and AliA (aliA) gene, partial cds.	Streptococcus pneumoniae	40,652	9-Sep-99
GB_EST9:AA073381	429	AA073381	mm83a12.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone IMAGE 535966 5' similar to gb:X03765 Mouse mRNA for cytoplasmic beta-actin (MOUSE);, mRNA sequence.	Mus musculus	46,341	15-Feb-97
GB_PR4:AC004874	95983	AC004874	Homo sapiens PAC clone DJ0744D13 from Tq11, complete sequence.	Homo sapiens	40,791	21-Nov-98
GB_EST14:AA418305	440	AA418305	zv96g05.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:767672 3', mRNA sequence.	Homo sapiens	40,093	16-OCT-1997
GB_BA2:U32763	12021	U32763	Haemophilus influenzae Rd section 78 of 163 of the complete genome.	Haemophilus influenzae Rd	37,715	29-MAY-1998
GB_BA1:MTCY10D7	39800	Z79700	Mycobacterium tuberculosis H37Rv complete genome; segment 44/162.	Mycobacterium tuberculosis	62,083	17-Jun-98
GB_BA1:MTCY10D7	39800	Z79700	Mycobacterium tuberculosis H37Rv complete genome; segment 44/162.	Mycobacterium tuberculosis	38,462	17-Jun-98
GB_GSS15:AAQ639831	649	AQ639831	927P1-20A2 TV 927P1 Trypanosoma brucei genomic clone 927P1-20A2, genomic survey sequence.	Trypanosoma brucei	36,124	8-Jul-99
GB_PR2:HS242N11	167514	AL023655	Human DNA sequence from clone 242N11 on chromosome 6p22.3-23. Contains ESTs, STSs, GSSs, genomic marker D6S285, and ca and gaa repeat polymorphisms, complete sequence.	Homo sapiens	33,866	23-Nov-99
GB_PR2:HS242N11	167514	AL023655	Human DNA sequence from clone 242N11 on chromosome 6p22.3-23. Contains ESTs, STSs, GSSs, genomic marker D6S285, and ca and gaa repeat polymorphisms, complete sequence.	Homo sapiens	36,342	23-Nov-99
GB_HTG1:CEY116A8_4110000	Z98858	Z98858	Caenorhabditis elegans chromosome IV clone Y116A8, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Caenorhabditis elegans	36,615	23-Jan-98
GB_IN1:CEY116A8B	29344	AL021469	Caenorhabditis elegans cosmid Y116A8B, complete sequence.	Caenorhabditis elegans	36,615	23-Nov-98
GB_HTG1:CEY116A8_4110000	Z98858	Z98858	Caenorhabditis elegans chromosome IV clone Y116A8, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Caenorhabditis elegans	36,615	23-Jan-98
GB_HTG5:AC007809	141122	AC007809	Drosophila melanogaster chromosome 3 clone BACR45M03 (D718) RPCI-98 45 M.3 map 88C-88C strain y; cn bw sp, *** SEQUENCING IN PROGRESS *** 98 unordered pieces.	Drosophila melanogaster	37,255	16-Nov-99
GB_IN2:CEL T02C5	36267	U55374	Caenorhabditis elegans cosmid T02C5.	Caenorhabditis elegans	41,834	08-OCT-1999
GB_IN1:CEZC101	26147	Z93395	Caenorhabditis elegans cosmid ZC101, complete sequence.	Caenorhabditis elegans	38,873	23-Jul-99
GB_EST16:C48630	360	C48630	C48630 Yuji Kohara unpublished cDNA: Strain N2 hermaphrodite embryo Caenorhabditis elegans cDNA clone yk459c11 5', mRNA sequence.	Caenorhabditis elegans	39,326	18-OCT-1999
GB_GSS12:AQ381854	597	AQ381854	RPCI11-137P5 TJ RPCI-11 Homo sapiens genomic clone RPCI-11-137P5, genomic survey sequence.	Homo sapiens	37,879	21-MAY-1999
GB_HTG3:AC010253	74835	AC010253	Homo sapiens chromosome 5 clone CIT-HSPC_432B14, *** SEQUENCING IN PROGRESS *** 34 unordered pieces.	Homo sapiens	35,082	15-Sep-99
GB_BA1:SC6G10	36734	AL049497	Streptomyces coelicolor cosmid 6G10.	Streptomyces coelicolor	40,047	24-MAR-1999
GB_BA1:MSGB1554CS	36548	L78814	Mycobacterium leprae cosmid B1554 DNA sequence.	Mycobacterium leprae	60,991	15-Jun-96
GB_BA1:MSGB1551CS	36548	L78813	Mycobacterium leprae cosmid B1551 DNA sequence.	Mycobacterium leprae	60,991	15-Jun-96
GB_HTG2:AC003118	98940	AC003118	Homo sapiens chromosome 1, *** SEQUENCING IN PROGRESS *** 3 unordered pieces.	Homo sapiens	34,008	24-Nov-97
GB_HTG2:AC003118	98940	AC003118	Homo sapiens chromosome 1, *** SEQUENCING IN PROGRESS *** 3 unordered pieces.	Homo sapiens	34,008	24-Nov-97
GB_PR3:HS395P12	147724	AL022310	Human DNA sequence from clone 395P12 on chromosome 1q24-25. Contains the TXGP1 gene for tax-transcriptionally activated glycoprotein 1 (34kD) (OX40 ligand, OX40L) and a GOT2 (Aspartate Aminotransferase, mitochondrial precursor, EC 2.6.1.1, Transaminase A, Glutamate Oxaloacetate Transaminase-2) pseudogene. Contains ESTs, STSs and GSSs, complete sequence.	Homo sapiens	34,008	23-Nov-99

TABLE 4: ALIGNMENT RESULTS

rx02146	750	GB_BA1:SC6G10	36734	AL049497	Streptomyces coelicolor cosmid 6G10.	Streptomyces coelicolor	36,167	24-MAR-1999
		GB_HTG6:AC008224	199774	AC008224	Drosophila melanogaster chromosome 3 clone BACR29J02 (D817) RPCI-98 29 J.2 map 83D-83D strain Y; cn bw sp. *** SEQUENCING IN PROGRESS***, 37 unordered pieces.	Drosophila melanogaster	36,437	24-Nov-99
		GB_GSS12:AQ407179	476	AQ407179	HS_5088_A2_G01_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=664 Col=2 Row=M, genomic survey sequence.	Homo sapiens	37,143	17-MAR-1999
rx02151	1311	GB_PR3:AC000074	44450	AC000074	Homo sapiens Chromosome 22q11.2 Cosmid Clone 20b In DSCR Region, complete sequence.	Homo sapiens	37,798	31-OCT-1998
		GB_PR2:HSAC002122	79931	AC002122	Human unknown clone GS293J04 from 5p15.2, complete sequence.	Homo sapiens	36,064	27-MAY-1997
		GB_PR2:HSAC002122	79931	AC002122	Human unknown clone GS293J04 from 5p15.2, complete sequence.	Homo sapiens	35,051	27-MAY-1997
rx02152	525	GB_PAT:AR009990	2793	AR009990	Sequence 1 from patent US 5756677.	Unknown.	44,311	04-DEC-1998
		GB_IN1:AF027735	3080	AF027735	Nephila clavipes minor ampullate silk protein MiSp1 mRNA, partial cds.	Nephila clavipes	44,311	23-Apr-98
		GB_PAT:IS5876	2793	IS5876	Sequence 1 from patent US 5733771.	Unknown.	44,311	01-DEC-1998
rx02163	876	GB_PR3:HSJ117516	96276	AL049538	Human DNA sequence from clone 117516 on chromosome 20. Contains the gene for Ras inhibitor JC265 (Ras association [RalGDS/AF-6] domain containing protein), ESTs, STSs, GSSs and two putative CpG islands, complete sequence.	Homo sapiens	36,385	23-Nov-99
		GB_GSS4:AQ690740	975	AQ690740	nbx0082N18f CUGI Rice BAC Library Oryza sativa genomic clone nbx0082N18f, genomic survey sequence.	Oryza sativa	39,474	1-Jul-99
rx02164	1698	GB_PH:SVWWBORF	6051	X72092	Streptomyces venezuelae bacteriophage VWB orfs.	Bacteriophage VWB	41,416	24-Aug-95
		GB_BA2:AF000579	8931	AF000579	Bordetella bronchiseptica LysR transcriptional activator homolog (bbuR), urease accessory protein D (ureD), urease structural subunit A (ureA), urease accessory protein J (ureJ), urease structural subunits B (ureB) and C (ureC), urease accessory proteins EF (ureEF) and G (ureG) genes, complete cds.	Bordetella bronchiseptica	38,005	23-Feb-98
		GB_PR3:HS431P23	147971	AL009178	Human DNA sequence from clone 431P23 on chromosome 6q27. Contains the first coding exon of the MLLT4 gene for myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 4 (AF-6, Afadin, MLLT-4, ALL-1 fusion partner), and a Serine Palmitoyltransferase 2 (EC 2.3.1.50, Long Chain Base Biosynthesis protein 2, LCB-2, SPT-2) pseudogene. Contains ESTs, STSs, GSSs, and a putative CpG island, complete sequence.	Homo sapiens	35,604	23-Nov-99
rx02165	735	GB_PR2:AB016897	331211	AB016897	Homo sapiens genomic DNA, chromosome 6q27, complete sequence.	Homo sapiens	38,314	20-Nov-99
		GB_IN2:AF079177	2159	AF079177	Theileria parva strain KNP2 p67 surface antigen (p67) gene, complete cds.	Theileria parva	39,207	09-MAR-1999
		GB_IN2:AF079176	2285	AF079176	Theileria parva strain Hluhuwe3 p67 surface antigen (p67) gene, complete cds.	Theileria parva	39,049	09-MAR-1999
		GB_EST27:AI397572	482	AI397572	NCS05G2T3 Subtracted Conidial Neurospora crassa cDNA clone SC5G2 5', mRNA sequence.	Neurospora crassa	38,936	8-Feb-99
rx02166	300	GB_BA1:MLCB1351	38936	Z95117	Mycobacterium leprae cosmid B1351.	Mycobacterium leprae	37,627	24-Jun-97
		GB_BA1:MLCB1351	38936	Z95117	Mycobacterium leprae cosmid B1351.	Mycobacterium leprae	40,741	24-Jun-97
rx02168	2937	GB_BA1:MAFASGEN	10520	X87822	B. ammoniagenes FAS gene.	Corynebacterium ammoniagenes	61,515	03-OCT-1996
		GB_BA1:MLCL458	43839	AL049478	Mycobacterium leprae cosmid L458.	Mycobacterium leprae	50,292	27-Aug-99
		GB_PR3:HSL19H1	40145	Z68164	Human DNA sequence from cosmid L19H1, Huntington's Disease Region, chromosome 4p16.3, containing multiple EST matches.	Homo sapiens	37,621	23-Nov-99
rx02169	969	GB_RO:MMA3CA212	4784	X94406	Mus musculus partial b3 gene for alpha3 subunit of L-type Ca2+ channel, exons 2-13.	Mus musculus	37,620	24-Nov-99
		GB_RO:RATCACH3B	2525	M88751	Rat calcium channel beta subunit-III mRNA, complete cds.	Rattus norvegicus	36,345	27-MAY-1994
		GB_RO:MMU20372	2469	U20372	Mus musculus voltage-dependent calcium channel beta-3 subunit (CCHB3) mRNA, complete cds.	Mus musculus	36,865	02-MAR-1996
rx02170	897	GB_HTG6:AC005497	212097	AC005497	Homo sapiens chromosome 17 clone RP11-952N18 map 17, *** SEQUENCING IN PROGRESS ***; 2 ordered pieces.	Homo sapiens	36,101	20-Nov-99
		GB_HTG6:AC005497	212097	AC005497	Homo sapiens chromosome 17 clone RP11-952N18 map 17, *** SEQUENCING IN PROGRESS ***; 2 ordered pieces.	Homo sapiens	37,570	20-Nov-99
		GB_PR2:AP000104	100000	AP000104	Homo sapiens genomic DNA of 21q22.1, GART and AML related, Q78C10-149C3 region, segment 7/20, complete sequence.	Homo sapiens	34,871	25-Sep-99
rx02172	462	GB_EST24:AI170522	367	AI170522	EST16450 Normalized rat lung, Bento Soares Rattus sp. cDNA clone RLUCO75 3' end, mRNA sequence.	Rattus sp.	38,904	20-Jan-99

TABLE 4: ALIGNMENT RESULTS

GB_GSS13:AQ430048	538	AQ430048	HS_5061_B2_G09_TTA RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=637 Col=18 Row=N, genomic survey sequence.	Homo sapiens	37,143	31-MAR-1999
GB_GSS4:AQ701704	589	AQ701704	HS_2130_A1_E03_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2130 Col=5 Row=I, genomic survey sequence.	Homo sapiens	35,280	7-Jul-99
GB_HTG4:AC009403	198495	AC009403	Homo sapiens clone DJ1015024, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Homo sapiens	37,154	28-OCT-1999
GB_HTG4:AC009403	198495	AC009403	Homo sapiens clone DJ1015024, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Homo sapiens	37,154	28-OCT-1999
GB_PR2:AP000094	100000	AP000094	Homo sapiens genomic DNA of 21q22.1, GART and AML related, B335D16-P10G11 region, segment 4/7, complete sequence.	Homo sapiens	38,536	25-Sep-99
rx02178						
GB_BA1:SC6C5	18160	AL034492	Streptomyces coelicolor cosmid 6C5.	Streptomyces coelicolor	53,635	14-DEC-1998
GB_HTG4:AC011121	172050	AC011121	Homo sapiens chromosome 11 clone 364_C_06 map 11, *** SEQUENCING IN PROGRESS ***; 9 ordered pieces.	Homo sapiens	37,170	14-OCT-1999
GB_HTG4:AC011121	172050	AC011121	Homo sapiens chromosome 11 clone 364_C_06 map 11, *** SEQUENCING IN PROGRESS ***; 9 ordered pieces.	Homo sapiens	37,170	14-OCT-1999
GB_PL1:STU76701	3049	U76701	Solanum tuberosum NADH nitrate reductase (SNR2) mRNA, complete cds.	Solanum tuberosum	38,846	18-Nov-96
GB_EST36:AI896605	633	AI896605	EST266048 tomato callus, TAMU Lycopersicon esculentum cDNA clone cLEC16C12, mRNA sequence.	Lycopersicon esculentum	39,671	27-Jul-99
GB_EST38:AW035744	620	AW035744	EST281898 tomato callus, TAMU Lycopersicon esculentum cDNA clone cLEC38M13, mRNA sequence.	Lycopersicon esculentum	40,555	15-Sep-99
rx02183						
GB_BA1:MTV043	68848	AL022004	Mycobacterium tuberculosis H37Rv complete genome; segment 40/162.	Mycobacterium tuberculosis	38,793	24-Jun-99
GB_BA1:MSG81935CS	40085	L04686	M. leprae genomic sequence, cosmid b1935.	Mycobacterium leprae	38,067	14-Jun-96
GB_BA1:MLCB57	38029	Z99494	Mycobacterium leprae cosmid B57.	Mycobacterium leprae	38,067	10-Feb-99
GB_BA2:CHSPAX03	307120	AJ248285	Pyrococcus abyssi complete genome; segment 3/6.	Pyrococcus abyssi	35,873	9-Sep-99
GB_HTG3:AC010677	116108	AC010677	Homo sapiens clone MS2304L04, *** SEQUENCING IN PROGRESS ***; 4 unordered pieces.	Homo sapiens	39,145	04-OCT-1999
GB_HTG3:AC010677	116108	AC010677	Homo sapiens clone MS2304L04, *** SEQUENCING IN PROGRESS ***; 4 unordered pieces.	Homo sapiens	39,145	04-OCT-1999
GB_EST33:AI779784	362	AI779784	EST260663 tomato susceptible, Cornell Lycopersicon esculentum cDNA clone cLES9K1, mRNA sequence.	Lycopersicon esculentum	41,011	29-Jun-99
GB_EST33:AI779784	362	AI779784	EST260663 tomato susceptible, Cornell Lycopersicon esculentum cDNA clone cLES9K1, mRNA sequence.	Lycopersicon esculentum	41,011	29-Jun-99
GB_PR3:AC005274	205150	AC005274	Homo sapiens chromosome 17, clone hRPK_1090_M_7, complete sequence.	Homo sapiens	36,337	29-Jul-98
GB_HTG2:AC007732	110348	AC007732	Homo sapiens chromosome 17 clone hRPC_1030_A_12 map 17, *** SEQUENCING IN PROGRESS ***; 7 unordered pieces.	Homo sapiens	35,924	5-Jun-99
GB_IN1:DMU11052	4871	U11052	Drosophila melanogaster peroxidase precursor mRNA, complete cds.	Drosophila melanogaster	39,873	26-Jan-95
GB_EST17:AA642571	323	AA642571	nc73c08.s1 NCL CGAP_P22 Homo sapiens cDNA clone IMAGE:1157966 3' similar to TR:G9372 G9372 UBQUITIN, mRNA sequence	Homo sapiens	49,485	27-OCT-1997
GB_EST27:AI425489	312	AI425489	my32e11.y1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA clone IMAGE:697580 5', mRNA sequence.	Mus musculus	40,514	09-MAR-1999
GB_EST23:AI154481	308	AI154481	ua03n07.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1345693 5', mRNA sequence.	Mus musculus	42,748	30-Sep-98
GB_PR2:AP000269	70932	AP000269	Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone:T293, complete sequence.	Homo sapiens	39,306	20-Nov-99
GB_PL1:D32140	1360	D32140	Cyanidioschyzon merolae DNA for actin, complete cds.	Cyanidioschyzon merolae	37,331	7-Feb-99
GB_PR2:AP000269	70932	AP000269	Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone:T293, complete sequence.	Homo sapiens	35,481	20-Nov-99

TABLE 4: ALIGNMENT RESULTS

rx02207	898	GB_EST21:AA982901	442	AA982901	vx59f07.r1 Stratiagene mouse macrophage (#937306) Mus musculus cDNA clone IMAGE:1279525 5' similar to gb.X59543_ma1RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M1 CHAIN (HUMAN); gb.K02927 Mouse ribonucleotide reductase subunit M1 mRNA, complete (MOUSE); mRNA sequence.	Mus musculus	42,141	27-MAY-1998
		GB_EST8:W85369	558	W85369	mi48h04.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:408343 5' similar to gb.K02927 Mouse ribonucleotide reductase subunit M1 mRNA, complete (MOUSE); mRNA sequence.	Mus musculus	39,964	12-Sep-96
rx02211		GB_PAT:E14824	2379	E14824	cDNA encoding M1 subunit of human ribonucleotide reductase.	Homo sapiens	36,812	28-Jul-99
rx02212	621	GB_HTG3:AC009518	241128	AC009518	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS *** 68 unordered pieces.	Homo sapiens	36,542	1-Sep-99
		GB_HTG3:AC009518	241128	AC009518	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS *** 68 unordered pieces.	Homo sapiens	36,542	1-Sep-99
		GB_HTG3:AC009518	241128	AC009518	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS *** 68 unordered pieces.	Homo sapiens	41,362	1-Sep-99
rx02216	329	GB_HTG2:AC007641	102402	AC007641	Mus musculus chromosome 10 clone 644_M_8 map 10, *** SEQUENCING IN PROGRESS *** 7 unordered pieces.	Mus musculus	31,307	22-MAY-1999
		GB_HTG2:AC007641	102402	AC007641	Mus musculus chromosome 10 clone 644_M_8 map 10, *** SEQUENCING IN PROGRESS *** 7 unordered pieces.	Mus musculus	31,307	22-MAY-1999
		GB_HTG2:AC007641	102402	AC007641	Mus musculus chromosome 10 clone 644_M_8 map 10, *** SEQUENCING IN PROGRESS *** 7 unordered pieces.	Mus musculus	32,198	22-MAY-1999
rx02217	786	GB_RO:RNU91516	1540	U91516	Rattus norvegicus oxytocin receptor (OTR) gene, promoter region.	Rattus norvegicus	36,256	25-Nov-97
		GB_RO:RNU91516	1540	U91516	Rattus norvegicus oxytocin receptor (OTR) gene, promoter region.	Rattus norvegicus	37,176	25-Nov-97
rx02218	390	GB_HTG3:AC008656	43033	AC008656	Homo sapiens chromosome 5 clone CIT9785KB_194J6, *** SEQUENCING IN PROGRESS *** 64 unordered pieces.	Homo sapiens	39,894	3-Aug-99
		GB_HTG3:AC008656	43033	AC008656	Homo sapiens chromosome 5 clone CIT9785KB_194J6, *** SEQUENCING IN PROGRESS *** 64 unordered pieces.	Homo sapiens	39,894	3-Aug-99
		GB_HTG3:AC008656	43033	AC008656	Homo sapiens chromosome 5 clone CIT9785KB_194J6, *** SEQUENCING IN PROGRESS *** 64 unordered pieces.	Homo sapiens	37,598	3-Aug-99
rx02219	1509	GB_PR3:HS475N16	113109	AL035587	Human DNA sequence from clone 475N16 on chromosome 6p12.3-21.2, complete sequence.	Homo sapiens	36,266	23-Nov-99
		GB_PR3:HS475N16	113109	AL035587	Human DNA sequence from clone 475N16 on chromosome 6p12.3-21.2, complete sequence.	Homo sapiens	35,255	23-Nov-99
		GB_EST10:AA142336	411	AA142336	ms07f02.r1 Stratiagene mouse skin (#937313) Mus musculus cDNA clone IMAGE:606267 5' similar to SW:RCA1_YEAST P40341 MITOCHONDRIAL RESPIRATORY CHAIN COMPLEXES ASSEMBLY PROTEIN RCA1, mRNA sequence.	Mus musculus	38,398	12-Feb-97
rx02221	1485	GB_PR3:AC002422	160091	AC002422	Human Chromosome X, complete sequence.	Homo sapiens	38,435	30-Jan-98
		GB_HTG2:AC006755	199917	AC006755	Caenorhabditis elegans clone Y40C5, *** SEQUENCING IN PROGRESS *** 1 unordered pieces.	Caenorhabditis elegans	39,229	23-Feb-99
		GB_HTG2:AC006755	199917	AC006755	Caenorhabditis elegans clone Y40C5, *** SEQUENCING IN PROGRESS *** 1 unordered pieces.	Caenorhabditis elegans	39,229	23-Feb-99
rx02223	601	GB_HTG6:AC008224	199774	AC008224	Drosophila melanogaster chromosome 3 clone BACR29J02 (D617) RPCI-98 29.J.2 map 83D-83D strain y, on bw sp, *** SEQUENCING IN PROGRESS *** 37 unordered pieces.	Drosophila melanogaster	35,333	24-Nov-99
		GB_HTG3:AC011511	158296	AC011511	Homo sapiens chromosome 19 clone CITB-H1_2369P2, *** SEQUENCING IN PROGRESS *** 57 unordered pieces.	Homo sapiens	35,593	07-OCT-1999
		GB_HTG3:AC011511	158296	AC011511	Homo sapiens chromosome 19 clone CITB-H1_2369P2, *** SEQUENCING IN PROGRESS *** 57 unordered pieces.	Homo sapiens	35,593	07-OCT-1999
rx02226	1156	GB_HTG2:AC006890	298195	AC006890	Caenorhabditis elegans clone Y67D8x, *** SEQUENCING IN PROGRESS *** 23 unordered pieces.	Caenorhabditis elegans	36,810	24-Feb-99
		GB_HTG2:AC006890	298195	AC006890	Caenorhabditis elegans clone Y67D8x, *** SEQUENCING IN PROGRESS *** 23 unordered pieces.	Caenorhabditis elegans	35,433	24-Feb-99
		GB_HTG2:AC006890	298195	AC006890	Caenorhabditis elegans clone Y67D8x, *** SEQUENCING IN PROGRESS *** 23 unordered pieces.	Caenorhabditis elegans	36,810	24-Feb-99



TABLE 4: ALIGNMENT RESULTS

rx02227	741	GB_EST5:H98835	440	H98835	yx14f12.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone IMAGE:261743 3', mRNA sequence.	Homo sapiens	38,182	15-DEC-1995
		GB_EST5:N25530	586	N25530	yx76c03.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone IMAGE:267652 3', mRNA sequence.	Homo sapiens	37,543	29-DEC-1995
		GB_BA1:CGUA43535	2531	U43535	Corynebacterium glutamicum multidrug resistance protein (cmr) gene, complete cds.	Corynebacterium glutamicum	38,812	9-Apr-97
rx02230	660	GB_EST16:AA570388	527	AA570388	nk62c08.s1 NCI_CGAP_Sch1 Homo sapiens cDNA clone IMAGE:1018094 3' similar to contains element MSR1 repetitive element 1; mRNA sequence.	Homo sapiens	41,096	9-Sep-97
		GB_PR2:HSUA43030	1539	U43030	Human cardiotrophin-1 (CTF1) mRNA, complete cds.	Homo sapiens	36,335	9-Jan-96
		GB_IN1:CELK08B5	35728	U41022	Caenorhabditis elegans cosmid K08B5.	Caenorhabditis elegans	35,387	30-Nov-95
rx02231	879	GB_PL2:ATAC006429	94818	AC006429	Arabidopsis thaliana chromosome II BAC F15K19 genomic sequence, complete sequence.	Arabidopsis thaliana	33,687	23-Apr-99
		GB_PR3:AC004554	195142	AC004554	Homo sapiens Xp22 BAC GSHB-590J6 (Genome Systems Human BAC library) complete sequence.	Homo sapiens	38,921	30-MAY-1998
		GB_HTG2:AC007853	116280	AC007853	Drosophila melanogaster chromosome 3 clone BACR03L02 (D766) RPCI-98 03.L.2 map 96B-96C strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 80 unordered pieces.	Drosophila melanogaster	36,237	2-Aug-99
rx02238	408	GB_BA1:MTCY21B4	39150	Z80108	Mycobacterium tuberculosis H37Rv complete genome; segment 62/162.	Mycobacterium tuberculosis	52,206	23-Jun-98
		GB_BA1:MTCY21B4	39150	Z80108	Mycobacterium tuberculosis H37Rv complete genome; segment 62/162.	Mycobacterium tuberculosis	36,935	23-Jun-98
		GB_VI:AF105451	801	AF105451	HIV-1 isolate A-DII-07 from Italy, envelope glycoprotein, C2-V5 region (env) gene, partial cds.	Human immunodeficiency virus type 1	40,864	25-Apr-99
rx02244	1656	GB_PR1:AB016195	10558	AB016195	Homo sapiens ELK1 pseudogene (ELK2) and immunoglobulin heavy chain gamma pseudogene (IGHGP).	Homo sapiens	38,929	9-Apr-99
		GB_PR3:HSN21F1	39212	Z94162	Human DNA sequence from cosmid N21F1 on chromosome 22 Contains exon trap and STS, complete sequence.	Homo sapiens	38,763	23-Nov-99
rx02254	800	GB_IN1:DMAAC001648	51989	AC001648	Drosophila melanogaster (P1 DS03431:D102) DNA sequence, complete sequence.	Drosophila melanogaster	36,884	22-Apr-97
		GB_GSS12:AQ360240	541	AQ360240	HS_5035_A2_E07_T7 RPCI11 Human Male BAC Library Homo sapiens genomic clone Plate=611 Col=14 Row=I, genomic survey sequence.	Homo sapiens	36,386	06-MAR-1999
		GB_GSS11:AQ258453	620	AQ258453	nbxb0020L15f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0020L15f, genomic survey sequence.	Oryza sativa	36,271	23-OCT-1998
		GB_GSS14:AQ577777	568	AQ577777	nbxb0091L17f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0091L17f, genomic survey sequence.	Oryza sativa	39,456	2-Jun-99
rx02255	1059	GB_PR4:AC006332	153477	AC006332	Homo sapiens clone NH0376O14, complete sequence.	Homo sapiens	35,019	11-Nov-99
		GB_PAT:E02669	1197	E02669	DNA encoding Bacillus sp. L-lactic acid dehydrogenase.	Bacillus sp.	39,719	29-Sep-97
		GB_PL1:YSCBO12	3201	D38310	Yeast BOI2 gene for Bcl2p.	Saccharomyces cerevisiae	35,769	8-Feb-99
rx02266	636	GB_BA1:REGIONB	4961	Z13995	N meningitidis lipA and lipB genes for LipA and LipB proteins.	Neisseria meningitidis	35,691	29-DEC-1993
		GB_HTG1:AP000568	136627	AP000568	Homo sapiens chromosome 21 clone B753B2 map 21q21.2, *** SEQUENCING IN PROGRESS ***. in unordered pieces.	Homo sapiens	37,500	07-OCT-1999
		GB_HTG1:AP000568	136627	AP000568	Homo sapiens chromosome 21 clone B753B2 map 21q21.2, *** SEQUENCING IN PROGRESS ***. in unordered pieces.	Homo sapiens	37,500	07-OCT-1999
rx02267	996	GB_PR4:AC007283	127361	AC007283	Homo sapiens clone NH0536I18, complete sequence.	Homo sapiens	37,155	28-Sep-99
		GB_IN1:CEC54C6	35500	Z77131	Caenorhabditis elegans cosmid C54C6, complete sequence.	Caenorhabditis elegans	38,280	23-Nov-98
		GB_HTG3:AC008905	129915	AC008905	Homo sapiens chromosome 5 clone CITB-H1_2259114, *** SEQUENCING IN PROGRESS ***. 40 unordered pieces.	Homo sapiens	35,895	3-Aug-99
rx02271	681	GB_HTG3:AC007441	219832	AC007441	Drosophila melanogaster chromosome 3 clone BACR10E03 (D690) RPCI-98 10.E.3 map 88A-88B strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 188 unordered pieces.	Drosophila melanogaster	33,284	08-OCT-1999
		GB_HTG3:AC007441	219832	AC007441	Drosophila melanogaster chromosome 3 clone BACR10E03 (D690) RPCI-98 10.E.3 map 88A-88B strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 188 unordered pieces.	Drosophila melanogaster	33,284	08-OCT-1999
		GB_HTG2:AC008029	123186	AC008029	Drosophila melanogaster chromosome 3 clone BACR01C11 (D819) RPCI-98 01.C.11 map 84D-84D strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***. 92 unordered pieces.	Drosophila melanogaster	34,315	2-Aug-99

TABLE 4: ALIGNMENT RESULTS

rx02279	1581	GB_IN1:AC003121 GB_IN1:AC003121 GB_BA1:MTCV39	69822 69822 38500	AC003121 AC003121 Z74025	Drosophila melanogaster (P1 DS00329 (D89)) DNA sequence, complete sequence. Drosophila melanogaster (P1 DS00329 (D89)) DNA sequence, complete sequence. Mycobacterium tuberculosis H37Rv complete genome; segment 89/162.	38,046 39,220 39,665	Drosophila melanogaster Drosophila melanogaster Mycobacterium tuberculosis	26-Nov-97 26-Nov-97 17-Jun-98
rx02280								
rx02286	672	GB_BA1:MTV025	121125	AL022121	Mycobacterium tuberculosis H37Rv complete genome; segment 155/162.	36,953	Mycobacterium tuberculosis	24-Jun-99
rx02287	675	GB_PL2:AC002130 GB_PL2:AC007259 GB_HTG3:AC009281 GB_HTG3:AC009281	114738 97146 221178 221178	AC002130 AC007259 AC009281 AC009281	The sequence of BAC F1N21 from Arabidopsis thaliana chromosome 1, complete sequence. Arabidopsis thaliana chromosome 1BAC T28P6 genomic sequence, complete sequence. Homo sapiens chromosome 15 clone 8_C_22 map 15. *** SEQUENCING IN PROGRESS ***. 49 unordered pieces. Homo sapiens chromosome 15 clone 8_C_22 map 15. *** SEQUENCING IN PROGRESS ***. 49 unordered pieces.	39,216 36,090 35,565	Arabidopsis thaliana Arabidopsis thaliana Homo sapiens	8-Jan-98 17-Aug-99 12-Aug-99
rx02294	498	GB_RO:MUSMURINC GB_EST19:AA813194 GB_EST19:AA813194	4597 398 398	M65736 AA813194 AA813194	Mouse murinoglobulin mRNA, complete cds. a180g12.s1 Soares_testis_NHT Homo sapiens cDNA clone 1377190 3', mRNA sequence. a180g12.s1 Soares_testis_NHT Homo sapiens cDNA clone 1377190 3', mRNA sequence.	35,565	Homo sapiens	12-Aug-99
rx02295	903	GB_PAT:189451 GB_VI:HCU33331 GB_PR3:HS550H1	18318 18535 108803	189451 U33331 AL035420	Sequence 6 from patent US 5721354. Human cytomegalovirus Toledo strain ULb' region. Human DNA sequence from clone 550H1 on chromosome 20q11.1-11.22 Contains a pseudogene similar to HIGH MOBILITY GROUP PROTEIN 2A, a novel mRNA, ESTs, STSs, GSSs and CpG Islands, complete sequence.	38,839 38,839 38,796	Unknown. human herpesvirus 5 Homo sapiens	10-Aug-98 27-Jan-96 23-Nov-99
rx02296	612	GB_BA2:AF065312 GB_PR3:HS550H1	1694 108803	AF065312 AL035420	Yersinia pestis hypothetical protein (yceG) gene, partial cds; thymidylate kinase (tmk) gene, complete cds; and putative DNA polymerase III delta' subunit (holB) gene, partial cds. Human DNA sequence from clone 550H1 on chromosome 20q11.1-11.22 Contains a pseudogene similar to HIGH MOBILITY GROUP PROTEIN 2A, a novel mRNA, ESTs, STSs, GSSs and CpG Islands, complete sequence.	41,351 37,919	Yersinia pestis Homo sapiens	16-Nov-99 23-Nov-99
rx02297	1260	GB_BA2:AF065312 GB_RO:AF007836 GB_IN2:EGU27015 GB_BA2:AF116184	1694 5655 2394 540	AF065312 AF007836 U27015 AF116184	Yersinia pestis hypothetical protein (yceG) gene, partial cds; thymidylate kinase (tmk) gene, complete cds; and putative DNA polymerase III delta' subunit (holB) gene, partial cds. Rattus norvegicus rab3 effector (RIM) mRNA, alternatively spliced, complete cds. Echinococcus granulosus 18S ribosomal RNA gene, complete sequence. Corynebacterium glutamicum L-aspartate-alpha-decarboxylase precursor (panD) gene, complete cds.	39,663 37,844 38,710 44,231	Yersinia pestis Rattus norvegicus Echinococcus granulosus Corynebacterium glutamicum	16-Nov-99 15-Aug-97 16-Jul-96 02-MAY-1999
rx02298	1782	GB_BA2:AF116184	540	AF116184	Corynebacterium glutamicum L-aspartate-alpha-decarboxylase precursor (panD) gene, complete cds.	42,007	Corynebacterium glutamicum	02-MAY-1999
rx02300	456	GB_BA2:AF116184	540	AF116184	Corynebacterium glutamicum L-aspartate-alpha-decarboxylase precursor (panD) gene, complete cds.	46,250	Corynebacterium glutamicum	02-MAY-1999
rx02301	840	GB_BA1:MTV004 GB_EST22:A1048692 GB_HTG3:AC008573	69350 172 205755	AL009198 A1048692 AC008573	Mycobacterium tuberculosis H37Rv complete genome; segment 144/162. ub30g04.11 Soares 2NbMT Mus musculus cDNA clone IMAGE:1379286 5', mRNA sequence. Homo sapiens chromosome 5 clone CIT-HSPC_551111, *** SEQUENCING IN PROGRESS ***. 95 unordered pieces.	39,602 43,284 35,115	Mycobacterium tuberculosis Mus musculus Homo sapiens	18-Jun-98 8-Jul-98 3-Aug-99
		GB_HTG3:AC008573	205755	AC008573	Homo sapiens chromosome 5 clone CIT-HSPC_551111, *** SEQUENCING IN PROGRESS ***. 95 unordered pieces.	35,115	Homo sapiens	3-Aug-99

TABLE 4: ALIGNMENT RESULTS

GB_HTG3:AC008573	205755	AC008573	Homo sapiens chromosome 5 clone CIT-HSPC_55111, *** SEQUENCING IN PROGRESS ***	95	Homo sapiens	36,527	3-Aug-99
GB_HTG6:AC008076	200000	AC008076	unordered pieces.				
GB_HTG3:AC008930	258026	AC008930	Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***	18	Homo sapiens	34,884	02-DEC-1999
GB_HTG3:AC008930	258026	AC008930	Homo sapiens chromosome 5 clone CITB-H1_2292M9, *** SEQUENCING IN PROGRESS ***	166	Homo sapiens	36,945	3-Aug-99
GB_HTG3:AC008930	258026	AC008930	unordered pieces.				
GB_HTG3:AC008930	258026	AC008930	Homo sapiens chromosome 5 clone CITB-H1_2292M9, *** SEQUENCING IN PROGRESS ***	166	Homo sapiens	36,945	3-Aug-99
GB_HTG3:AC008930	258026	AC008930	unordered pieces.				
GB_PAT:A69720	53789	A69720	Sequence 3 from Patent WO9807868.		unidentified	36,915	07-MAY-1999
GB_BA1:AMM223012	53784	AJ223012	Amycolatopsis mediterranei genes encoding rifamycin polyketide synthases, ORFs 1 to 5.		Amycolatopsis mediterranei	36,915	9-Feb-98
GB_BA2:AF040570	78199	AF040570	Amycolatopsis mediterranei rifamycin biosynthetic gene cluster.		Amycolatopsis mediterranei	36,915	5-Feb-98
GB_HTG3:AC009340	110415	AC009340	Drosophila melanogaster chromosome 2 clone BACR04E19 (D1026) RPCI-98 04 E. 19 map 34A-34E strain y, cn bw sp, *** SEQUENCING IN PROGRESS ***	103	Drosophila melanogaster	36,066	27-Aug-99
GB_HTG3:AC009340	110415	AC009340	Drosophila melanogaster chromosome 2 clone BACR04E19 (D1026) RPCI-98 04 E. 19 map 34A-34E strain y, cn bw sp, *** SEQUENCING IN PROGRESS ***	103	Drosophila melanogaster	36,066	27-Aug-99
GB_IN2:AC005656	48443	AC005656	Drosophila melanogaster, chromosome 2R, region 34B6-34C2, P1 clone DS08787, complete sequence.		Drosophila melanogaster	46,154	6-Feb-99
GB_HTG5:AC011247	206436	AC011247	Homo sapiens clone NH0541E12, WORKING DRAFT SEQUENCE, 1 unordered pieces.		Homo sapiens	36,201	13-Nov-99
GB_HTG5:AC011247	206436	AC011247	Homo sapiens clone NH0541E12, WORKING DRAFT SEQUENCE, 1 unordered pieces.		Homo sapiens	35,036	13-Nov-99
GB_HTG3:AC011152	142616	AC011152	Homo sapiens clone 7_H_4, LOW-PASS SEQUENCE SAMPLING.		Homo sapiens	35,548	01-OCT-1999
GB_BA2:AF110185	20302	AF110185	Burkholderia pseudomallei strain 1026b DnbB (dnhB), general secretory pathway protein D (gspD), general secretory pathway protein E (gspE), general secretory pathway protein F (gspF), GspC (gspC), general secretory pathway protein G (gspG), general secretory pathway protein H (gspH), general secretory pathway protein I (gspI), general secretory pathway protein J (gspJ), general secretory pathway protein K (gspK), general secretory pathway protein L (gspL), general secretory pathway protein M (gspM), and general secretory pathway protein N (gspN) genes, complete cds; and unknown genes.		Burkholderia pseudomallei	39,670	2-Aug-99
GB_BA1:PSEAMNH	5215	D90216	P. chlororaphis genes for amidase (EC 3.5.1.4) and for nitrile hydratase (EC 4.2.1.84).		Pseudomonas chlororaphis	51,254	7-Feb-99
GB_PAT:E12519	4775	E12519	Nucleotide sequence of Rhodococcus rhodochrous genomic DNA region containing amidase and nitrilhydratase genes.		Rhodococcus rhodochrous	51,646	24-Jun-98
GB_BA1:CGPYC	3728	Y09548	Corynebacterium glutamicum pyc gene.		Corynebacterium glutamicum	100,000	08-MAY-1998
GB_PR4:AC006079	178109	AC006079	Homo sapiens chromosome 17, clone hRPK 855_D_21, complete sequence.		Homo sapiens	37,807	12-DEC-1998
GB_GSS8:AQ036832	441	AQ036832	CIT-HSP-2334L1, TF CIT-HSP Homo sapiens genomic clone 2334L1, genomic survey sequence.		Homo sapiens	42,359	11-Jul-98
GB_PL1:YSKGA11	4159	M68870	Kluyveromyces fragilis transcriptional activator (GAL11) gene, complete cds.		Kluyveromyces fragilis	40,252	14-MAY-1993
GB_PR2:HSAY18950	584	Y18950	Homo sapiens partial gene for caspase-9, intronic sequence (584 bp).		Homo sapiens	40,529	08-OCT-1999
GB_PL1:YSKGA11	4159	M68870	Kluyveromyces fragilis transcriptional activator (GAL11) gene, complete cds.		Kluyveromyces fragilis	38,679	14-MAY-1993
GB_BA1:CGU35023	3195	U35023	Corynebacterium glutamicum thiosulfate sulfurtransferase (thiR) gene, partial cds, acyl CoA carboxylase (accBC) gene, complete cds.		Corynebacterium glutamicum	35,548	16-Jan-97
GB_PL2:AF114171	183990	AF114171	Sorghum bicolor BAC clone 25.M18, complete sequence.		Sorghum bicolor	41,414	25-Apr-99
GB_PR4:AC006324	157310	AC006324	Homo sapiens clone DJ1164F05, complete sequence.		Homo sapiens	41,786	11-Nov-99
GB_BA1:CGU35023	3195	U35023	Corynebacterium glutamicum thiosulfate sulfurtransferase (thiR) gene, partial cds, acyl CoA carboxylase (accBC) gene, complete cds.		Corynebacterium glutamicum	100,000	16-Jan-97

**TABLE 4: ALIGNMENT RESULTS**

GB_BA1:MTCY22G8	22550	Z95585	Mycobacterium tuberculosis H37Rv complete genome; segment 49/162.	38,072	17-Jun-98
GB_PR2:HSAY18950	584	Y18950	Homo sapiens partial gene for caspase-9, intronic sequence (584 bp).	42,949	08-OCT-1999
GB_PR2:HS149A16	173354	AL021937	Human DNA sequence from clone 149A16 on chromosome 22q12-13. Contains an IGLC (immunoglobulin Lambda Chain C) pseudogene, the RFPL3 and RFPL3S genes for Ret finger protein-like 3 and Ret finger protein-like 3 antisense respectively, a gene for a novel Immunoglobulin Lambda Chain V family protein, a novel gene for a protein similar to mouse RGDS (RALGDS, RALGEF, Guanine Nucleotide Dissociation Stimulator A) and rabbit oncogene RSC, a novel gene for the human ortholog of worm F16A11.2 and bacterial and archaea-bacterial predicted proteins, a novel gene for a protein similar to BPI (Bacterial Permeability-Increasing Protein) and rabbit LBP (Liposaccharide-Binding Protein), and the 5' part of a novel gene. Contains ESTs, STSs, GSSs, genomic marker D22S1175, a ca repeat polymorphism and putative CpG islands, complete sequence.	37,376	23-Nov-99
GB_HTG3:AC009726	177618	AC009726	Homo sapiens chromosome 18 clone 263_O_14 map 18, *** SEQUENCING IN PROGRESS ***; 12 Homo sapiens unordered pieces.	37,037	29-Aug-99
GB_HTG3:AC009726	177618	AC009726	Homo sapiens chromosome 18 clone 263_O_14 map 18, *** SEQUENCING IN PROGRESS ***; 12 Homo sapiens unordered pieces.	37,037	29-Aug-99
GB_VI:AF063866	236120	AF063866	Melanoplus sanguinipes entomopoxvirus, complete genome.	34,480	22-DEC-1998
GB_HTG4:AC011089	171283	AC011089	Homo sapiens chromosome 2 clone 303_K_20 map 2, *** SEQUENCING IN PROGRESS ***; 25 ordered pieces.	37,757	14-OCT-1999
GB_HTG4:AC011089	171283	AC011089	Homo sapiens chromosome 2 clone 303_K_20 map 2, *** SEQUENCING IN PROGRESS ***; 25 ordered pieces.	37,757	14-OCT-1999
GB_GSS14:AQ580594	303	AQ580594	RPC1-11-452K5; TJ RPC1-11 Homo sapiens genomic clone RPC1-11-452K5, genomic survey sequence.	38,305	7-Jun-99
GB_GSS12:AQ393454	561	AQ393454	CITBI-E1-2556E3; TR CITBI-E1 Homo sapiens genomic clone 2556E3, genomic survey sequence.	37,882	06-MAR-1999
GB_GSS14:AQ507468	627	AQ507468	RPC1-11-298O19; TJ RPC1-11 Homo sapiens genomic clone RPC1-11-298O19, genomic survey sequence.	38,480	29-Apr-99
GB_PR3:AC005549	147416	AC005549	Homo sapiens chromosome 17, clone hRPK.215_E_13, complete sequence.	39,909	22-Sep-98
GB_PR3:AC005549	147416	AC005549	Homo sapiens chromosome 17, clone hRPK.215_E_13, complete sequence.	33,784	22-Sep-98
GB_GSS8:B92789	742	B92789	CIT-HSP-2164J8; TR CIT-HSP Homo sapiens genomic clone 2164J8, genomic survey sequence.	39,409	25-Jun-98
GB_EST19:AA738949	396	AA738949	w68d12.1r1 Striatogene mouse skin (#937313) Mus musculus cDNA clone IMAGE:1227575 5', mRNA sequence.	36,023	14-Jan-98
GB_GSS14:AQ526017	500	AQ526017	HS_5329_A2_F02_T7A RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=905 Col=4 Row=K, genomic survey sequence.	38,537	11-MAY-1999
GB_PR3:U82695	167460	U82695	Homo sapiens cosmid LM1937 from Xq28.	38,421	3-Jan-98
GB_BA1:MTCY1687	43430	Z81331	Mycobacterium tuberculosis H37Rv complete genome; segment 123/162.	39,087	17-Jun-98
GB_BA1:SRMSIK	2384	Y08921	S. reticulii gene encoding Msk protein and orf1.	58,356	21-MAR-1997
GB_BA1:MSGV414A	40121	AD000007	Mycobacterium tuberculosis sequence from clone y414a.	58,423	03-DEC-1996
GB_BA1:BSUB0021	215534	Z99124	Bacillus subtilis complete genome (section 21 of 21); from 3999281 to 4214814.	36,829	26-Nov-97
GB_PR3:HSDA22P16	103367	AL049641	Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone DA22P16, complete sequence.	36,430	23-Nov-99
GB_BA1:BSUB0021	215534	Z99124	Bacillus subtilis complete genome (section 21 of 21); from 3999281 to 4214814.	37,440	26-Nov-97
GB_EST21:AA970555	420	AA970555	oo94h05 s1 NCI CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1573881 3' similar to gb:X61970 PROTEASOME ZETA CHAIN (HUMAN); mRNA sequence.	36,905	20-MAY-1998

במסגרת פיקוח השופט, נא להגיש את כל המסמכים הנדרשים לתיק, כולל תצהירי אמת, עדויות, חוות דעת, וכל מסמך אחר הקשור למסגרת הפיקוח.

TABLE 4: ALIGNMENT RESULTS

GB_GSS4:AQ739589	909	AQ739589	HS_5381_B2_G06_T7A RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=957 Col=12 Row=N, genomic survey sequence.	Homo sapiens	44,912	16-Jul-99
GB_IN1:CEY53H1C	37004	AL117201	Caenorhabditis elegans cosmid Y53H1C, complete sequence.	Caenorhabditis elegans	38,436	19-Nov-99
GB_IN2:CEUK04F10	35413	AF039719	Caenorhabditis elegans cosmid K04F10.	Caenorhabditis elegans	36,891	26-MAY-1999
GB_PL1:AB011477	78181	AB011477	Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: MHK7, complete sequence.	Arabidopsis thaliana	37,095	20-Nov-99
GB_PL1:AB011477	78181	AB011477	Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: MHK7, complete sequence.	Arabidopsis thaliana	36,794	20-Nov-99
GB_EST10:AA178985	357	AA178985	zpl2g08.s1 Stralagene fetal retina 937202 Homo sapiens cDNA clone IMAGE:609278 3', mRNA sequence.	Homo sapiens	39,875	31-DEC-1996
GB_GSS4:AQ710468	555	AQ710468	HS_5336_A2_B09_T7A RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=912 Col=18 Row=C, genomic survey sequence.	Homo sapiens	39,820	13-Jul-99
GB_EST37:AI986900	380	AI986900	rs19a09.y1 Sommer Pristionchus Pristionchus pacificus cDNA clone IMAGE: 5' similar to WP:C08A1.1 CE02114 TRANSITIONAL ENDOPLASMIC RETICULUM ATPASE HOMOLOG 1.; mRNA sequence.	Pristionchus pacificus	42,105	1-Sep-99
GB_BA2:SCF76	18292	AL121600	Streptomyces coelicolor cosmid F76.	Streptomyces coelicolor A3(2)	40,195	29-Sep-99
GB_SY:SCU53587	4546	U53587	Artificial Corynebacterium glutamicum IS1207-derived transposon transposase genes, complete cds, and 3'5'-aminoglycoside phosphotransferase (aphA-3) gene, complete cds.	synthetic construct	36,755	06-MAY-1996
GB_PAT:E16763	2517	E16763	gDNA encoding aspartate transferase (AAT).	Corynebacterium glutamicum	38,687	28-Jul-99
rxa02367	732					
rxa02368						
GB_PL2:AC010924	80442	AC010924	Arabidopsis thaliana chromosome 1 BAC T24D18 sequence, complete sequence.	Arabidopsis thaliana	39,261	6-Nov-99
GB_VIHE1CG	152261	X14112	Herpes simplex virus (HSV) type 1 complete genome.	human herpesvirus 1	39,856	17-Apr-97
GB_VIHS1ULR	108360	D10879	Herpes simplex virus type 1 long unique region UL.	human herpesvirus 1	39,856	3-Feb-99
GB_BA1:CGPROAGEN	1783	X82929	C glutamicum proA gene.	Corynebacterium glutamicum	98,974	23-Jan-97
GB_BA1:CGPROAGEN	1783	X82929	C glutamicum proA gene.	Corynebacterium glutamicum	37,156	23-Jan-97
GB_GSS10:AQ215523	445	AQ215523	HS_2259_B2_F03_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2259 Col=6 Row=L, genomic survey sequence.	Homo sapiens	40,674	19-Sep-98
GB_EST35:AI810729	450	AI810729	tt86h09.x1 NCL_CGAP_P28 Homo sapiens cDNA clone IMAGE:2248481 3', mRNA sequence.	Homo sapiens	41,568	7-Jul-99
GB_EST22:AI048725	347	AI048725	ub31d01.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1379329 5', mRNA sequence.	Mus musculus	39,412	8-Jul-98
GB_EST15:AA466288	431	AA466288	vr34c01.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA clone IMAGE:888864 5', mRNA sequence.	Mus musculus	39,336	11-Jun-97
GB_PR3:AC004687	175120	AC004687	Homo sapiens chromosome 17, clone hRPC.1171_1_10, complete sequence.	Homo sapiens	34,518	26-Jun-98
GB_PR3:AC004687	175120	AC004687	Homo sapiens chromosome 17, clone hRPC.1171_1_10, complete sequence.	Homo sapiens	38,551	26-Jun-98
GB_GSS15:AQ604975	536	AQ604975	HS_2135_B2_G04_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2135 Col=8 Row=N, genomic survey sequence.	Homo sapiens	39,321	10-Jun-99
GB_BA2:SC51A	42527	AL121596	Streptomyces coelicolor cosmid 51A.	Streptomyces coelicolor A3(2)	38,804	28-Sep-99
GB_BA2:AF073776	5138	AF073776	Pseudomonas aeruginosa MexZ (mexZ), complete cds; and mexGH operon, complete sequence.	Pseudomonas aeruginosa	37,245	19-OCT-1998
GB_BA1:AB015853	5461	AB015853	Pseudomonas aeruginosa gene for MexX and MexY, complete cds.	Pseudomonas aeruginosa	40,641	13-Nov-98
GB_BA1:RCU57682	86896	U57682	Rhodobacter capsulatus cosmid 143-147, complete sequence.	Rhodobacter capsulatus	44,246	7-Feb-97
GB_PL2:T2K10	88037	AC005966	Arabidopsis thaliana chromosome 1 BAC T2K10 sequence, complete sequence.	Arabidopsis thaliana	42,455	10-Feb-99
GB_PL2:T24H24	88848	AF075598	Arabidopsis thaliana BAC T24H24.	Arabidopsis thaliana	34,826	3-Aug-98
GB_BA1:SC5F2A	40105	AL049587	Streptomyces coelicolor cosmid 5F2A.	Streptomyces coelicolor	51,688	24-MAY-1999
GB_BA1:CGBE1PGEN	2339	X93514	C glutamicum betP gene.	Corynebacterium glutamicum	48,877	8-Sep-97



TABLE 4: ALIGNMENT RESULTS

rx02427	510	GB_EST27:AI438343	610	AI438343	SWOVAFCAP31F09SK Onchocerca volvulus adult female cDNA (SAW98MLW-OvAF) Onchocerca volvulus cDNA clone SWOVAFCAP31F09 5' mRNA sequence.	Onchocerca volvulus	41,244	09-MAR-1999
		GB_BA1:XANHHPA1A	1824	M99173	Xanthomonas campestris HrpA1 gene, complete cds.	Xanthomonas campestris	37,520	14-Sep-93
		GB_OV:CCA245635	2455	AJ245635	Cyprinus carpio IL-1 gene for interleukin-1-beta.	Cyprinus carpio	35,614	5-Aug-99
		GB_OV:AB010701	1213	AB010701	Cyprinus carpio mRNA for interleukin-1 beta, complete cds.	Cyprinus carpio	36,821	29-Jan-98
		GB_OV:CCA245635	2455	AJ245635	Cyprinus carpio IL-1 gene for interleukin-1-beta.	Cyprinus carpio	35,060	5-Aug-99
rx02428	1257	GB_IN1:CEY40B1B	29313	AL032636	Caenorhabditis elegans cosmid Y40B1B, complete sequence.	Caenorhabditis elegans	37,068	12-Nov-99
		GB_GSS12:AQ374513	640	AQ374513	RPC111-145N15. TJ RPC1-11 Homo sapiens genomic clone RPC1-11-145N15, genomic survey sequence.	Homo sapiens	35,636	20-MAY-1999
rx02430	309	GB_IN1:CEY40B1B	29313	AL032636	Caenorhabditis elegans cosmid Y40B1B, complete sequence.	Caenorhabditis elegans	32,800	12-Nov-99
		GB_BA1:MTCY1A11	30850	Z78020	Mycobacterium tuberculosis H37Rv complete genome; segment 83/162.	Mycobacterium tuberculosis	64,610	17-Jun-98
		GB_PR4:AC004998	135572	AC004998	Homo sapiens clone DJ164D05, complete sequence.	Homo sapiens	41,118	17-Jun-99
		GB_PR2:AC002091	161799	AC002091	Genomic sequence from Human 17, complete sequence.	Homo sapiens	36,393	9-Sep-97
rx02433	723	GB_HTG5:AC009754	212978	AC009754	Homo sapiens chromosome 15 clone RP11-519C12, WORKING DRAFT SEQUENCE, 16 ordered pieces.	Homo sapiens	33,008	17-Nov-99
		GB_HTG5:AC009754	212978	AC009754	Homo sapiens chromosome 15 clone RP11-519C12, WORKING DRAFT SEQUENCE, 16 ordered pieces.	Homo sapiens	39,716	17-Nov-99
rx02437	933	GB_PR4:HUAAC003108	164564	AC003108	Human Chromosome 16 BAC clone CIT987SK-327024, complete sequence.	Homo sapiens	35,180	23-Nov-99
		GB_IN2:AC005650	60019	AC005650	Drosophila melanogaster, chromosome 2R, region 58B4-58B7, P1 clone DS02885, complete sequence.	Drosophila melanogaster	40,397	30-Jan-99
		GB_IN2:AC005650	60019	AC005650	Drosophila melanogaster, chromosome 2R, region 58B4-58B7, P1 clone DS02885, complete sequence.	Drosophila melanogaster	37,719	30-Jan-99
		GB_RO:AF081193	2250	AF081193	Mus musculus calcium and DAG-regulated guanine nucleotide exchange factor   mRNA, complete cds.	Mus musculus	37,835	28-Nov-98
rx02443	1077	GB_HTG4:AC009370	59409	AC009370	Drosophila melanogaster chromosome 3L75C1 clone RPC198-35F4, *** SEQUENCING IN PROGRESS ***; 40 unordered pieces.	Drosophila melanogaster	35,460	16-OCT-1999
		GB_HTG4:AC009370	59409	AC009370	Drosophila melanogaster chromosome 3L75C1 clone RPC198-35F4, *** SEQUENCING IN PROGRESS ***; 40 unordered pieces.	Drosophila melanogaster	35,460	16-OCT-1999
		GB_EST9:AA081445	311	AA081445	zn18a01.r1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone IMAGE:547752 5', mRNA sequence.	Homo sapiens	36,334	21-OCT-1996
rx02444	1401	GB_BA2:SCF76	18292	AL121600	Streptomyces coelicolor cosmid F76.	Streptomyces coelicolor A3(2)	39,226	29-Sep-99
		GB_PR2:HS1022J11	137658	AL049765	Human DNA sequence from clone 1022J11 on chromosome 20q13.13-13.2, complete sequence.	Homo sapiens	35,444	23-Nov-99
		GB_HTG1:AC002345	132645	AC002345	Homo sapiens chromosome 17 clone 20D5, *** SEQUENCING IN PROGRESS ***; 10 unordered pieces.	Homo sapiens	35,106	25-Aug-97
rx02452	303	GB_GSS14:AQ578499	825	AQ578499	nbxb0084D17f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0084D17f, genomic survey sequence.	Oryza sativa	39,535	2-Jun-99
		GB_GSS14:AQ578499	825	AQ578499	nbxb0084D17f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0084D17f, genomic survey sequence.	Oryza sativa	36,213	2-Jun-99
rx02454	1365	GB_HTG3:AC007810	140175	AC007810	Drosophila melanogaster chromosome 3 clone BACR14A01 (D720) RPC1-98 14.A.1 map 90C-90C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 89 unordered pieces.	Drosophila melanogaster	33,931	17-Sep-99
		GB_HTG3:AC007810	140175	AC007810	Drosophila melanogaster chromosome 3 clone BACR14A01 (D720) RPC1-98 14.A.1 map 90C-90C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 89 unordered pieces.	Drosophila melanogaster	33,931	17-Sep-99
		GB_PR3:AC005339	32360	AC005339	Homo sapiens chromosome 19, cosmid R33729, complete sequence.	Homo sapiens	38,131	30-Jul-98
rx02457	1233	GB_HTG3:AC009346	105005	AC009346	Drosophila melanogaster chromosome 3 clone BACR03P13 (D672) RPC1-98 03.P.13 map 83A-83B strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 83 unordered pieces.	Drosophila melanogaster	36,829	27-Aug-99
		GB_EST20:AA820427	436	AA820427	LD24042.5prime LD Drosophila melanogaster embryo pOT2 Drosophila melanogaster cDNA clone LD24042 5prime, mRNA sequence.	Drosophila melanogaster	38,018	25-Feb-99

TABLE 4: ALIGNMENT RESULTS

GB_HTG3:AC009346	105005	AC009346	Drosophila melanogaster chromosome 3 clone BACR03P13 (D672) RPCI-98 03 P.13 map 83A-83B strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 83 unordered pieces.	Drosophila melanogaster	36,829	27-Aug-99
GB_BA2:AF114233	1852	AF114233	Corynebacterium glutamicum 5-enolpyruvylshikimate 3-phosphate synthase (aroA) gene, complete cds.	Corynebacterium glutamicum	99,248	7-Feb-99
GB_HTG1:CEY45F10_2 110000	Z93245	Z93245	Caenorhabditis elegans chromosome IV clone Y45F10, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	35,621	17-Sep-97
GB_HTG1:CEY45F10_2 110000	Z93245	Z93245	Caenorhabditis elegans chromosome IV clone Y45F10, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	35,621	17-Sep-97
GB_EST10:AA153371	590	AA153371	ms02a12.r1 Stratagene mouse embryonic carcinoma (#937317) Mus musculus cDNA clone IMAGE:605758 5' similar to gb:M24194 GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE PROTEIN (HUMAN); gb:X75313 M.musculus (MOUSE); mRNA sequence.	Mus musculus	39,203	11-Feb-97
GB_EST10:AA153371	590	AA153371	ms02a12.r1 Stratagene mouse embryonic carcinoma (#937317) Mus musculus cDNA clone IMAGE:605758 5' similar to gb:M24194 GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE PROTEIN (HUMAN); gb:X75313 M.musculus (MOUSE); mRNA sequence.	Mus musculus	41,920	11-Feb-97
GB_HTG4:AC008754	86446	AC008754	Homo sapiens chromosome 19 clone CITB-E1_3023J11, *** SEQUENCING IN PROGRESS ***; 73 unordered pieces.	Homo sapiens	36,064	31-OCT-1999
GB_HTG4:AC008754	86446	AC008754	Homo sapiens chromosome 19 clone CITB-E1_3023J11, *** SEQUENCING IN PROGRESS ***; 73 unordered pieces.	Homo sapiens	36,064	31-OCT-1999
GB_HTG4:AC008754	86446	AC008754	Homo sapiens chromosome 19 clone CITB-E1_3023J11, *** SEQUENCING IN PROGRESS ***; 73 unordered pieces.	Homo sapiens	37,589	31-OCT-1999
GB_PR3:AC004960	143834	AC004960	Homo sapiens PAC clone DJ1098B01 from 7q11.23-q21, complete sequence.	Homo sapiens	35,010	5-Nov-98
GB_GSS9:AQ122158	428	AQ122158	HS_3083_A1_E10_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3083 Col=19 Row=M, genomic survey sequence.	Homo sapiens	39,720	22-Sep-98
GB_BA2:AF053227	840	AF053227	Photobacterium leiognathi probable flavin reductase (luxG) gene, complete cds.	Photobacterium leiognathi	37,572	12-Jun-98
GB_EST21:AA942401	825	AA942401	LD26583.5prime LD Drosophila melanogaster embryo pOT2 Drosophila melanogaster cDNA clone LD26583 5prime, mRNA sequence.	Drosophila melanogaster	37,705	25-Nov-98
GB_GSS12:AQ400564	485	AQ400564	HS_5064_A1_G09_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=640 Col=17 Row=M, genomic survey sequence.	Homo sapiens	38,043	13-MAR-1999
GB_GSS12:AQ400564	485	AQ400564	HS_5064_A1_G09_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=640 Col=17 Row=M, genomic survey sequence.	Homo sapiens	37,647	13-MAR-1999
GB_HTG2:AC005718	149592	AC005718	Drosophila melanogaster chromosome 2 clone DS02336 (D440) map 60C8-60D2 strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 68 unordered pieces.	Drosophila melanogaster	40,525	30-Jul-99
GB_HTG2:AC005718	149592	AC005718	Drosophila melanogaster chromosome 2 clone DS02336 (D440) map 60C8-60D2 strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 68 unordered pieces.	Drosophila melanogaster	40,525	30-Jul-99
GB_PR3:AC004217	103319	AC004217	Homo sapiens 12q24.1 PAC RPCI3-521E19 (Roswell Park Cancer Institute Human PAC library) complete sequence.	Homo sapiens	37,389	2-Jun-98
GB_GSS12:AQ413964	550	AQ413964	RPCI-11-207C17-TV RPCI-11 Homo sapiens genomic clone RPCI-11-207C17, genomic survey sequence.	Homo sapiens	37,870	23-MAR-1999
GB_IN2:CELR09H3	10687	U58740	Caenorhabditis elegans cosmid R09H3.	Caenorhabditis elegans	37,408	26-MAY-1999
GB_GSS13:AQ488397	734	AQ488397	RPCI-11-243H14-TV RPCI-11 Homo sapiens genomic clone RPCI-11-243H14, genomic survey sequence.	Homo sapiens	37,569	24-Apr-99
GB_HTG1:CEY7A5	235922	AL021576	Caenorhabditis elegans chromosome X clone Y7A5, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	38,561	19-Aug-99
GB_HTG1:CEY7A5	235922	AL021576	Caenorhabditis elegans chromosome X clone Y7A5, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	38,561	19-Aug-99
GB_IN1:CEY7A5A	68270	AL034489	Caenorhabditis elegans cosmid Y7A5A, complete sequence.	Caenorhabditis elegans	38,081	18-DEC-1998



**TABLE 4: ALIGNMENT RESULTS**

rx02475	1278	GB_BA1:AB009078	2686	AB009078	Brevibacterium saccharolyticum gene for L-2,3-butanediol dehydrogenase, complete cds.	Brevibacterium saccharolyticum	99,217	13-Feb-99
		GB_PH:AF074945	15644	AF074945	Mycoplasma arthritis bacteriophage MAV1, complete genome.	Mycoplasma arthritis bacteriophage MAV1	36,876	27-Jul-99
		GB_BA1:MTV008	63033	AL021246	Mycobacterium tuberculosis H37Rv complete genome; segment 108/162.	Mycobacterium tuberculosis	37,520	17-Jun-98
rx02478	1338	GB_HTG1:AC002419	128340	AC002419	Homo sapiens chromosome X clone bWXD40, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	34,700	12-Aug-97
		GB_PR3:AC004073	79612	AC004073	Human Chromosome X, complete sequence.	Homo sapiens	34,700	29-Jan-98
		GB_HTG1:AC002419	128340	AC002419	Homo sapiens chromosome X clone bWXD40, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	34,700	12-Aug-97
rx02482	933	GB_OV:AF077329	1645	AF077329	Coturnix coturnix pro-alpha2(I) collagen mRNA, partial cds.	Coturnix coturnix	40,202	3-Feb-99
		GB_PL2:SFU59150	2854	U59150	Sartorya fumigata nuclear protein AICB5p (AICB5p) mRNA, complete cds.	Aspergillus fumigatus	36,789	1-Jan-98
		GB_BA2:AE000106	12554	AE000106	Rhizobium sp. NGR234 plasmid pNGR234a, section 43 of 46 of the complete plasmid sequence.	Rhizobium sp. NGR234	36,630	12-DEC-1997
rx02483	936	GB_HTG1:AC002345	132645	AC002345	Homo sapiens chromosome 17 clone 20D5, *** SEQUENCING IN PROGRESS***; 10 unordered pieces.	Homo sapiens	37,942	25-Aug-97
		GB_HTG1:AC002345	132645	AC002345	Homo sapiens chromosome 17 clone 20D5, *** SEQUENCING IN PROGRESS ***; 10 unordered pieces.	Homo sapiens	37,942	25-Aug-97
		GB_PR4:AF118569	24070	AF118569	Homo sapiens angiotensin I converting enzyme precursor (DCP1) gene, alternative splice products, complete cds.	Homo sapiens	37,352	28-OCT-1999
rx02484	624	GB_BA1:RLSPRLVCP	3696	Y09534	R. leguminosarum Symbiosis Plasmid DNA, rlvCP gene.	Rhizobium leguminosarum	38,003	26-Feb-97
		GB_BA2:RLU23040	3931	U23040	Rhizobium leguminosarum bv. viciae putative glycerol-3-phosphate transport protein (ugpC) gene, partial cds, and chemoreceptor protein (mcpA), putative 2-hydroxychromene-2-carboxylate isomerase, and putative alcohol dehydrogenase genes, complete cds.	Rhizobium leguminosarum bv. viciae	38,003	26-Aug-96
		GB_BA1:NGORBKGMF	1580	L07845	Neisseria gonorrhoeae ribokinase (rbk) gene, 3' end; ADP-L-glycero-D-mannoheptose epimerase (gme) gene, complete cds.	Neisseria gonorrhoeae	46,939	11-OCT-1995
rx02486	759	GB_IN1:CELM02D8	40876	U41034	Caenorhabditis elegans cosmid M02D8.	Caenorhabditis elegans	38,358	30-Nov-95
		GB_IN2:CELC36C5	41596	AF016444	Caenorhabditis elegans cosmid C36C5.	Caenorhabditis elegans	38,482	08-OCT-1999
		GB_HTG3:AC011467	188118	AC011467	Homo sapiens chromosome 19 clone CIT-HSPC_457E21, *** SEQUENCING IN PROGRESS ***; 59 unordered pieces.	Homo sapiens	37,954	07-OCT-1999
rx02488	392	GB_HTG2:AC006765	274498	AC006765	Caenorhabditis elegans clone Y43H11, *** SEQUENCING IN PROGRESS ***; 7 unordered pieces.	Caenorhabditis elegans	38,961	23-Feb-99
		GB_HTG2:AC006765	274498	AC006765	Caenorhabditis elegans clone Y43H11, *** SEQUENCING IN PROGRESS ***; 7 unordered pieces.	Caenorhabditis elegans	38,961	23-Feb-99
		GB_HTG4:AC010035	173152	AC010035	Drosophila melanogaster chromosome 3L/74B2 clone RPCO98-6H1, *** SEQUENCING IN PROGRESS ***; 60 unordered pieces.	Drosophila melanogaster	38,010	16-OCT-1999
rx02489	724	GB_GSS8:AQ050210	796	AQ050210	rbxb0003d008r CUGI Rice BAC Library Oryza sativa genomic clone rbxb0003H15r, genomic survey sequence.	Oryza sativa	40,201	24-MAR-1999
		GB_PL1:D87819	2036	D87819	Oryza sativa mRNA for sucrose transporter, complete cds.	Oryza sativa	39,356	26-DEC-1997
		GB_STS:G53565	627	G53565	SHGC-83995 Human Homo sapiens STS genomic, sequence tagged site.	Homo sapiens	40,237	25-Jun-99
rx02495	2691	GB_BA1:IMTCY19H9	20679	Z83857	Mycobacterium tuberculosis H37Rv complete genome; segment 129/162.	Mycobacterium tuberculosis	40,853	18-Jun-98
		GB_BA1:IMTCY19H9	20679	Z83857	Mycobacterium tuberculosis H37Rv complete genome; segment 129/162.	Mycobacterium tuberculosis	40,358	18-Jun-98
		GB_HTG1:HSDJ90108	114599	AL078461	Homo sapiens chromosome 20 clone RP5-901O8 map q11.1-11.23, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	34,059	25-Nov-99
rx02496	930	GB_EST36:AV201062	373	AV201062	AV201062 Yuiji Kohara unpublished cDNA Caenorhabditis elegans cDNA clone yk25011 5', mRNA sequence.	Caenorhabditis elegans	41,287	26-Jul-99
		GB_IN1:CELC10F3	40172	AF022968	Caenorhabditis elegans cosmid C10F3.	Caenorhabditis elegans	35,217	10-Sep-97

TABLE 4: ALIGNMENT RESULTS

rx02498	927	GB_EST29:AI612578	357	AI612578	TENG0372 T. Cruzi epimastigote normalised cDNA Library Trypanosoma cruzi cDNA clone n715.r 5'. mRNA sequence.	Trypanosoma cruzi	39,474	7-Jul-99
		GB_BA2:CGU31225	1817	U31225	Corynebacterium glutamicum L-proline:NADP+ 5-oxidoeductase (proC) gene, complete cds.	Corynebacterium glutamicum	92,580	2-Aug-96
		GB_PR2:AP000197	100000	AP000197	Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone B355D16-T1073, segment 3/8, complete sequence.	Homo sapiens	36,344	20-Nov-99
		GB_PR2:AP000093	100000	AP000093	Homo sapiens genomic DNA of 21q22.1, GART and AML related, B335D16-P10G11 region, segment 3/7, complete sequence.	Homo sapiens	36,344	25-Sep-99
rx02500	222	GB_BA1:MTCY20G9	37218	Z77162	Mycobacterium tuberculosis H37Rv complete genome; segment 25/162.	Mycobacterium tuberculosis	58,257	17-Jun-98
		GB_BA1:U00018	42991	U00018	Mycobacterium leprae cosmid B2168.	Mycobacterium leprae	44,776	01-MAR-1994
		GB_BA1:SCE68	38084	AL079345	Streptomyces coelicolor cosmid E68.	Streptomyces coelicolor A3(2)	37,624	16-Jul-99
rx02505	294	GB_RO:MM2B4J	289	X00621	Mouse hybridoma 2B4 gene fragment (J-region) for T-cell receptor.	Mus musculus	36,364	10-Nov-95
		GB_RO:MUSTCBJB1	2276	K02802	Mouse T-cell receptor germ-line beta-chain J-beta-2 gene cluster, including J-beta-[2.1.2.2.3.2.4.2.5.2.psi.2.6] genes.	Mus musculus	32,862	27-Apr-93
		GB_RO:MMAE000665	199101	AE000665	Mus musculus TCR beta locus from bases 501860 to 700960 (section 3 of 3) of the complete sequence.	Mus musculus	36,879	4-Sep-97
rx02506								
rx02510	769	GB_OM:PIGMNTNADP	1595	M86719	Pig mitochondrial NADP-isocitrate dehydrogenase mRNA, 3' end.	Sus scrofa	36,963	12-Jun-93
		GB_EST17:AA629042	359	AA629042	zu77008.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:744015 3', mRNA sequence.	Homo sapiens	37,184	16-OCT-1997
		GB_EST20:AA837524	528	AA837524	oe31007.s1 NCL CGAP_P125 Homo sapiens cDNA clone IMAGE:1410157 similar to gb:X69391 60S RIBOSOMAL PROTEIN L6 (HUMAN); mRNA sequence.	Homo sapiens	40,039	26-Feb-98
rx02514	857	GB_PR2:CNS01DS9	175148	AL121694	Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-701B16 of RPCI-11 library from chromosome 14 of Homo sapiens (Human), complete sequence.	Homo sapiens	35,185	8-Nov-99
		GB_PR4:AC006222	165643	AC006222	Homo sapiens, clone hRPK.12_A_1, complete sequence.	Homo sapiens	33,411	15-Jan-99
		GB_PR4:AC006222	165643	AC006222	Homo sapiens, clone hRPK.12_A_1, complete sequence.	Homo sapiens	35,629	15-Jan-99
rx02518	534	GB_BA1:MTV007	32806	AL021184	Mycobacterium tuberculosis H37Rv complete genome; segment 64/162.	Mycobacterium tuberculosis	57,965	17-Jun-98
		GB_EST16:AA576027	231	AA576027	nm57e05.s1 NCL CGAP_Br3 Homo sapiens cDNA clone IMAGE:1072352 3' similar to TR:G1001455 G1001455 HYPOTHETICAL 141.7 KD PROTEIN. ., mRNA sequence.	Homo sapiens	47,674	9-Sep-97
		GB_BA2:AE000226	10466	AE000226	Escherichia coli K-12 MG1655 section 116 of 400 of the complete genome.	Escherichia coli	38,447	12-Nov-98
rx02519	1752	GB_IN2:DMNRG2	8574	AF050085	Drosophila melanogaster neuroglian (nrg) gene, exons 3-6, 7a, 7b and alternatively spliced products, complete cds.	Drosophila melanogaster	38,498	5-Aug-98
		GB_IN2:AC004322	32480	AC004322	Drosophila melanogaster DNA sequence (P1 DS01982 (D216)), complete sequence.	Drosophila melanogaster	36,126	29-Aug-98
		GB_HTG2:AC008188	129128	AC008188	Drosophila melanogaster chromosome 2 clone BACR08118 (D660) RPCI-98 08.118 map 56A2-56B1 strain Y; cn bw sp. *** SEQUENCING IN PROGRESS *** 72 unordered pieces.	Drosophila melanogaster	35,660	2-Aug-99
rx02520	807	GB_GSS3:B68449	228	B68449	CIT-HSP-2025P8.TF CIT-HSP Homo sapiens genomic clone 2025P8, genomic survey sequence.	Homo sapiens	42,544	21-Jun-98
		GB_HTG3:AC010642	45982	AC010642	Homo sapiens chromosome 19 clone LLNL-R_245B6, *** SEQUENCING IN PROGRESS *** 26 unordered pieces.	Homo sapiens	34,543	16-Sep-99
		GB_HTG3:AC010642	45982	AC010642	Homo sapiens chromosome 19 clone LLNL-R_245B6, *** SEQUENCING IN PROGRESS *** 26 unordered pieces.	Homo sapiens	34,543	16-Sep-99
rx02521	1566	GB_BA1:SCC22	22115	AL096839	Streptomyces coelicolor cosmid C22.	Streptomyces coelicolor	40,885	12-Jul-99
		GB_BA1:BSUB0017	217420	Z99120	Bacillus subtilis complete genome (section 17 of 21): from 3197001 to 3414420.	Bacillus subtilis	38,441	26-Nov-97
		GB_BA1:MLCL536	36224	Z99125	Mycobacterium leprae cosmid L536.	Mycobacterium leprae	37,233	04-DEC-1998
rx02524	813	GB_IN1:CELF42H10	28687	L08403	C. elegans cosmid F42H10.	Caenorhabditis elegans	32,878	22-OCT-1993
		GB_IN1:CEU02289	4824	U02289	Caenorhabditis elegans Bristol N2 GTPase-activating protein (CEGAP) mRNA, partial cds.	Caenorhabditis elegans	34,457	11-Jun-94

**TABLE 4: ALIGNMENT RESULTS**

rx02525	501	GB_IN1:CELF42H10	28687	L08403	C. elegans cosmid F42H10.	Caenorhabditis elegans	34,243	22-OCT-1993
		GB_EST36:AU076280	416	AU076280	Cu076280 Rice kernel shoot Oryza sativa cDNA clone S10896_11A, mRNA sequence.	Oryza sativa	42,671	27-Jul-99
		GB_EST1:D34202	355	D34202	CELK042ESR Yujii Kohara unpublished cDNA Caenorhabditis elegans cDNA clone yk42a5 3', mRNA sequence.	Caenorhabditis elegans	42,553	8-Aug-94
		GB_EST1:D34202	355	D34202	CELK042ESR Yujii Kohara unpublished cDNA Caenorhabditis elegans cDNA clone yk42a5 3', mRNA sequence.	Caenorhabditis elegans	37,892	8-Aug-94
rx02534	927	GB_EST15:AA490533	426	AA490533	aa51n04.r1 NCI CGAP_GCB1 Homo sapiens cDNA clone IMAGE:824503 5', mRNA sequence.	Homo sapiens	41,176	15-Aug-97
		GB_PL2:ATFCA3	200252	Z97338	Arabidopsis thaliana DNA chromosome 4, ESSA I FCA contig fragment No. 3.	Arabidopsis thaliana	35,568	27-Aug-99
		GB_PL2:ATFCA3	200252	Z97338	Arabidopsis thaliana DNA chromosome 4, ESSA I FCA contig fragment No. 3.	Arabidopsis thaliana	37,376	27-Aug-99
rx02537	585	GB_HTG4:AC009732	165576	AC009732	Drosophila melanogaster chromosome 2 clone BACR05E17 (D1059) RPCI-98 05.E.17 map 57F-57F strain y; on bw sp, *** SEQUENCING IN PROGRESS ***; 83 unordered pieces.	Drosophila melanogaster	30,727	26-OCT-1999
		GB_HTG4:AC009732	165576	AC009732	Drosophila melanogaster chromosome 2 clone BACR05E17 (D1059) RPCI-98 05.E.17 map 57F-57F strain y; on bw sp, *** SEQUENCING IN PROGRESS ***; 83 unordered pieces.	Drosophila melanogaster	30,727	26-OCT-1999
		GB_HTG4:AC009732	165576	AC009732	Drosophila melanogaster chromosome 2 clone BACR05E17 (D1059) RPCI-98 05.E.17 map 57F-57F strain y; on bw sp, *** SEQUENCING IN PROGRESS ***; 83 unordered pieces.	Drosophila melanogaster	36,460	26-OCT-1999
rx02538	795	GB_GSS15:AQ606090	658	AQ606090	HS_5382_A1_E12_TTA RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=968 Col=23 Row=I, genomic survey sequence.	Homo sapiens	39,369	10-Jun-99
		GB_BA1:AXACE	5323	X94981	A.xylimum aceB, aceC, aceD, and aceE genes.	Acetobacter xylinus	37,009	12-Feb-97
		GB_BA1:AXACE	5323	X94981	A.xylimum aceB, aceC, aceD, and aceE genes.	Acetobacter xylinus	37,879	12-Feb-97
rx02540	561	GB_BA1:MTY13E10	35019	Z95324	Mycobacterium tuberculosis H37Rv complete genome; segment 18/162.	Mycobacterium tuberculosis	51,002	17-Jun-98
		GB_BA1:MTDNAGRP	4644	X58406	M.tuberculosis dnaK, grpE, and dnaJ genes.	Mycobacterium tuberculosis	51,002	21-OCT-1992
		GB_BA1:AB007847	2400	AB007847	Rhodococcus rhodochrous gene for 3-ketosteroid-delta1-dehydrogenase, complete cds.	Rhodococcus rhodochrous	53,723	13-Feb-99
rx02544	1521	GB_PAT:E14041	2776	E14041	gDNA encoding laminaripentaose forming enzyme (LPHase).	Streptomyces sp.	47,351	28-Jul-99
		GB_HTG3:AC010592	45971	AC010592	Homo sapiens chromosome 5 clone CIT-HSPC_459H20, *** SEQUENCING IN PROGRESS ***; 38 unordered pieces.	Homo sapiens	37,689	16-Sep-99
		GB_HTG3:AC010592	45971	AC010592	Homo sapiens chromosome 5 clone CIT-HSPC_459H20, *** SEQUENCING IN PROGRESS ***; 38 unordered pieces.	Homo sapiens	37,689	16-Sep-99
rx02545	681	GB_HTG3:AC009016	128095	AC009016	Homo sapiens chromosome 5 clone P1_889E7, *** SEQUENCING IN PROGRESS ***; 67 unordered pieces.	Homo sapiens	34,763	3-Aug-99
		GB_HTG3:AC009016	128095	AC009016	Homo sapiens chromosome 5 clone P1_889E7, *** SEQUENCING IN PROGRESS ***; 67 unordered pieces.	Homo sapiens	34,763	3-Aug-99
		GB_HTG3:AC009016	72817	AC009016	Homo sapiens chromosome 5 clone P1_1352A1, *** SEQUENCING IN PROGRESS ***; 19 unordered pieces.	Homo sapiens	34,763	3-Aug-99
rx02546	1227	GB_HTG3:AC006519	134365	AC006519	Homo sapiens clone RPCI5-951N9, *** SEQUENCING IN PROGRESS ***; 41 unordered pieces.	Homo sapiens	32,868	16-Sep-99
		GB_BA1:D90738	17528	D90738	Escherichia coli genomic DNA (23.0 - 23.4 min).	Escherichia coli	38,674	7-Feb-99
		GB_HTG3:AC006519	134365	AC006519	Homo sapiens clone RPCI5-951N9, *** SEQUENCING IN PROGRESS ***; 41 unordered pieces.	Homo sapiens	32,868	16-Sep-99
rx02549	2826	GB_BA1:CGBPHI16	962	Y12472	C.glutamicum DNA, attachment site bacteriophage Phi-16.	Corynebacterium glutamicum	41,702	05-MAR-1999
		GB_BA2:AE001095	10592	AE001095	Archaeoglobus fulgidus section 12 of 172 of the complete genome.	Archaeoglobus fulgidus	36,817	15-DEC-1997
		GB_EST20:AA887411	587	AA887411	gb.X69908.ma1 ATP SYNTHASE LIPID-BINDING PROTEIN P2 PRECURSOR (HUMAN); mRNA sequence.	Homo sapiens	41,738	9-Jun-98
rx02552	918	GB_GSS9:AQ129371	377	AQ129371	HS_3045_A2_B07 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3045 Col=14 Row=C, genomic survey sequence.	Homo sapiens	42,857	23-Sep-98
		GB_GSS14:AQ566979	600	AQ566979	HS_2105_A2_D12 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2105 Col=24 Row=G, genomic survey sequence.	Homo sapiens	35,443	29-MAY-1999

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

rx02554	747	GB_IN1:MEPAP	4636	X54422	M.edulis gene for polyphenolic adhesive protein.	Mytilus edulis	35,619	27-MAY-1992
		GB_PR4:AC004882	171878	AC004882	Homo sapiens PAC clone DJ076B20 from 22, complete sequence.	Homo sapiens	37,346	22-OCT-1999
		GB_PR4:AC004882	171878	AC004882	Homo sapiens PAC clone DJ076B20 from 22, complete sequence.	Homo sapiens	38,440	22-OCT-1999
		GB_BA2:U32830	16388	U32830	Haemophilus influenzae Rd section 145 of 163 of the complete genome.	Haemophilus influenzae Rd	33,514	29-MAY-1998
rx02555	726	GB_GSS1:FR0020618	466	AL013501	F.rubripes GSS sequence, clone 042H13bD8, genomic survey sequence.	Fugu rubripes	40,215	10-DEC-1997
		GB_GSS1:FR0020576	555	AL013459	F.rubripes GSS sequence, clone 042H13aE5, genomic survey sequence.	Fugu rubripes	35,379	10-DEC-1997
		GB_GSS1:FR0020585	431	AL013468	F.rubripes GSS sequence, clone 042H13aH5, genomic survey sequence.	Fugu rubripes	39,718	10-DEC-1997
rx02564	1125	GB_PR3:AC004976	127425	AC004976	Homo sapiens PAC clone DJ1143H19 from 7p14-p15, complete sequence.	Homo sapiens	37,364	5-Nov-98
		GB_PR2:HUAC002038	161973	AC002038	Homo sapiens chromosome 2 clone 101B6 map 2p11, complete sequence.	Homo sapiens	36,404	30-Jun-97
		GB_HTG3:AC010791	111643	AC010791	Homo sapiens chromosome 17 clone 6_M_14 map 17, *** SEQUENCING IN PROGRESS *** , 8 unordered pieces.	Homo sapiens	34,293	22-Sep-99
rx02568	1753	GB_BA1:FT16SRNAA	1521	Z21931	F.tularensis 16S rRNA.	Francisella tularensis	39,065	5-Jun-97
		GB_BA1:FT16SRNAB	1520	Z21932	F.tularensis 16S rRNA.	Francisella tularensis	37,807	16-DEC-1997
		GB_BA1:FP16SRNAA	1518	Z21933	F.philomiragia 16S rRNA.	Francisella philomiragia	36,380	5-Jun-97
rx02569	873	GB_GSS13:AAQ469518	553	AAQ469518	CITB1E1-2596J20, TR CITB1E1 Homo sapiens genomic clone 2596J20, genomic survey sequence.	Homo sapiens	39,201	23-Apr-99
		GB_GSS12:AAQ43903	495	AAQ43903	HS_5058_B1_F07_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=634 Col=13 Row=L, genomic survey sequence.	Homo sapiens	36,049	13-MAR-1999
rx02570	765	GB_PR2:AF004715	2889	AF004715	Homo sapiens jerky gene product homolog mRNA, complete cds.	Homo sapiens	35,689	6-Aug-97
		GB_BA1:CGU43535	2531	U43535	Corynebacterium glutamicum multidrug resistance protein (cmr) gene, complete cds.	Corynebacterium glutamicum	38,165	9-Apr-97
		GB_EST33:AI776398	499	AI776398	EST257498 tomato resistant, Cornell Lycopersicon esculentum cDNA clone cLER1810, mRNA sequence.	Lycopersicon esculentum	42,084	29-Jun-99
rx02573		GB_EST23:AI119807	431	AI119807	uc21a06.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone IMAGE:1398610 5', mRNA sequence.	Mus musculus	40,281	2-Sep-98
rx02575	1258	GB_BA1:SCGD3	33779	AL096822	Streptomyces coelicolor cosmid GD3.	Streptomyces coelicolor	39,742	8-Jul-99
		GB_RO:AC002108	41125	AC002108	Genomic sequence from Mouse 4, complete sequence.	Mus musculus	36,100	29-MAY-1997
		GB_PR1:HUMIDS	36845	L35485	Homo sapiens iduronate sulphate sulphatase (IDS) gene, complete cds.	Homo sapiens	34,183	16-Aug-94
rx02576	1545	GB_GSS15:AAQ618143	582	AAQ618143	HS_5168_B1_H08_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=744 Col=15 Row=P, genomic survey sequence.	Homo sapiens	34,874	15-Jun-99
		GB_OV:XLFB1A1	2172	X64759	X.laevis mRNA for transcription factor (clone XLFB1a1).	Xenopus laevis	38,589	26-Jan-93
		GB_PR4:AC005039	151316	AC005039	Homo sapiens clone NH0512E16, complete sequence.	Homo sapiens	33,575	14-Jan-99
rx02577	1008	GB_PL2:AF034976	7130	AF034976	Playella littoralis ribosomal protein S14 (rps14) gene, partial cds; ATPase subunit 8 (atp8) gene, complete cds; rRNA-Ser gene, complete sequence; ribosomal protein S10 (rps10) and ribosomal protein L31 (rpl31) genes, complete cds; rRNA-Ser gene, complete sequence; T7-like RNA polymerase (rpoX) gene, complete cds; and unknown genes, mitochondrial genes for mitochondrial products.	Mitochondrion Playella littoralis	34,690	22-Jun-98
		GB_EST30:AI655188	456	AI655188	wb67g02.x1 NCI_CGAP_G06 Homo sapiens cDNA clone IMAGE:2310770 3', mRNA sequence.	Homo sapiens	39,198	04-MAY-1999
		GB_PL2:AF034976	7130	AF034976	Playella littoralis ribosomal protein S14 (rps14) gene, partial cds; ATPase subunit 8 (atp8) gene, complete cds; rRNA-Ser gene, complete sequence; ribosomal protein S10 (rps10) and ribosomal protein L31 (rpl31) genes, complete cds; rRNA-Ser gene, complete sequence; T7-like RNA polymerase (rpoX) gene, complete cds; and unknown genes, mitochondrial genes for mitochondrial products.	Mitochondrion Playella littoralis	39,959	22-Jun-98
rx02584	474	GB_PR2:AP000191	100000	AP000191	Homo sapiens genomic DNA, chromosome 21q22.1, D21S228-AML region, clone Q78C10-f32E9, segment 18/21, complete sequence.	Homo sapiens	37,199	20-Nov-99
		GB_EST13:AA349881	293	AA349881	EST56832 Infant brain Homo sapiens cDNA 5' end similar to EST containing Alu repeat, mRNA sequence.	Homo sapiens	42,466	21-Apr-97

TABLE 4: ALIGNMENT RESULTS

rx02585	1104	GB_EST23:AI078215	404	AI078215	oz12h12.x1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:1675175 3' similar to contains Alu repetitive element contains element THR repetitive element ;, mRNA sequence.	Homo sapiens	48,780	29-Sep-98
		GB_PR2:AP000115	100000	AP000115	Homo sapiens genomic DNA of 21q22.1, GART and AML related, Q78C10-149C3 region, segment 18/20, complete sequence.	Homo sapiens	35,502	25-Sep-99
		GB_PR2:AP000191	100000	AP000191	Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone Q78C10-f32E9, segment 18/21, complete sequence.	Homo sapiens	35,502	20-Nov-99
rx02588	807	GB_PR2:AP000047	50188	AP000047	Homo sapiens genomic DNA, chromosome 21q22.1, segment 18/28, complete sequence.	Homo sapiens	35,502	20-Nov-99
		GB_BA1:MTCY8D5	39730	Z92689	Mycobacterium tuberculosis H37Rv complete genome; segment 12/162.	Mycobacterium tuberculosis	38,354	18-Jun-98
		GB_BA1:MTCY349	43523	Z83018	Mycobacterium tuberculosis H37Rv complete genome; segment 131/162.	Mycobacterium tuberculosis	39,949	17-Jun-98
rx02591	1953	GB_HTG2:AC007879	158841	AC007879	Homo sapiens clone NH0309L06, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	36,835	31-Jul-99
		GB_BA1:MTCY8D5	39730	Z92669	Mycobacterium tuberculosis H37Rv complete genome; segment 12/162.	Mycobacterium tuberculosis	65,398	18-Jun-98
		GB_BA1:MLCL622	42498	Z95398	Mycobacterium leprae cosmid L622.	Mycobacterium leprae	63,875	24-Jun-97
		GB_BA1:CHBATP2OP	5477	L08777	Chlorobium limicola atp2 operon.	Chlorobium limicola	37,979	26-Apr-93
rx02593	720	GB_GSS1:CNS0056G	994	AL057090	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR11M23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	Drosophila melanogaster	27,206	3-Jun-99
		GB_GSS1:CNS0056G	994	AL057090	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR11M23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	Drosophila melanogaster	30,270	3-Jun-99
rx02598	714	GB_IN1:PFMAL3P5	86829	AL034556	Plasmodium falciparum MAL3P5, complete sequence.	Plasmodium falciparum	40,057	04-OCT-1999
		GB_PL1:AB024026	23026	AB024026	Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone: K15O15, complete sequence.	Arabidopsis thaliana	38,865	20-Nov-99
		GB_BA2:AE000930	15553	AE000930	Methanobacterium thermoautotrophicum from bases 1592014 to 1607566 (section 136 of 148) of the complete genome.	Methanobacterium thermoautotrophicum	36,804	15-Nov-97
rx02600	1521	GB_IN2:AC004301	68620	AC004301	Drosophila melanogaster DNA sequence (P1 DS07134 (D192)), complete sequence.	Drosophila melanogaster	36,063	29-MAY-1998
		GB_PL2:AC007504	125021	AC007504	Arabidopsis thaliana chromosome I BAC F13F21 genomic sequence, complete sequence.	Arabidopsis thaliana	37,074	9-Jul-99
		GB_PH:AF125520	61670	AF125520	Bacteriophage 933W, complete genome.	Bacteriophage 933W	35,396	16-Apr-99
rx02601	2112	GB_BA1:MTV026	23740	AL022076	Mycobacterium tuberculosis H37Rv complete genome; segment 157/162.	Mycobacterium tuberculosis	38,676	24-Jun-99
		GB_BA2:SSU73128	2331	U73128	Sphingomonas sp. A8AN3 catechol 2,3-dioxygenase gene, complete cds and 2-hydroxymuconic semialdehyde hydrolase and 2-hydroxymuconic semialdehyde dehydrogenase genes, partial cds.	Sphingomonas sp. A8AN3	40,020	1-Feb-99
		GB_BA2:SSU73128	2331	U73128	Sphingomonas sp. A8AN3 catechol 2,3-dioxygenase gene, complete cds and 2-hydroxymuconic semialdehyde hydrolase and 2-hydroxymuconic semialdehyde dehydrogenase genes, partial cds.	Sphingomonas sp. A8AN3	40,101	1-Feb-99
rx02602	627	GB_PR3:AF006751	3106	AF006751	Homo sapiens ES/130 mRNA, complete cds.	Homo sapiens	37,582	10-Jul-98
		GB_PR3:AF006751	3106	AF006751	Homo sapiens ES/130 mRNA, complete cds.	Homo sapiens	38,462	10-Jul-98
		GB_GSS9:AQ090116	404	AQ090116	HS_3000_B1_E09_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plates=3000 Col=17 Row=J, genomic survey sequence.	Homo sapiens	39,481	26-Aug-98
rx02604	2199	GB_PL1:CREHIH3G	4358	L41841	Chlamydomonas reinhardtii histone H3, histone H4, histone H2B, and histone H2A genes, complete cds.	Chlamydomonas reinhardtii	37,500	16-MAY-1996
		GB_GSS13:AQ455681	621	AQ455681	HS_5088_B1_E01_TTA RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=644 Col=1 Row=J, genomic survey sequence.	Homo sapiens	36,246	21-Apr-99
		GB_GSS8:AQ829471	392	AQ829471	HS_5442_A2_F11_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1018 Col=22 Row=K, genomic survey sequence.	Homo sapiens	42,347	27-Aug-99
rx02606	1494	GB_PAT:A26027	2547	A26027	C.melassecola gene for extracellular antigen PS1.	Corynebacterium melassecola	42,857	2-Apr-95
		GB_BA1:CGCOP1G	2547	X66078	C.glutamicum cop1 gene for PS1.	Corynebacterium glutamicum	42,857	30-Jun-93

TABLE 4: ALIGNMENT RESULTS

rx02609	372	GB_PR2:AP000134	100000	AP000134	Homo sapiens genomic DNA of 21q22.1, GART and AML, #3D11-11988 region, segment 9/10, complete sequence.	Homo sapiens	36,145	25-Sep-99
		GB_HTG3:AC009212	125452	AC009212	Drosophila melanogaster chromosome 3 clone BACR01A18 (D669) RPCI-98 01 A.18 map 82E-82F strain y; cn bw sp, *** SEQUENCING IN PROGRESS***, 119 unordered pieces.	Drosophila melanogaster	35,753	23-Aug-99
		GB_HTG3:AC009212	125452	AC009212	Drosophila melanogaster chromosome 3 clone BACR01A18 (D669) RPCI-98 01 A.18 map 82E-82F strain y; cn bw sp, *** SEQUENCING IN PROGRESS***, 119 unordered pieces.	Drosophila melanogaster	35,753	23-Aug-99
		GB_HTG2:AC007589	134659	AC007589	Drosophila melanogaster chromosome 3 clone BACR20D10 (D667) RPCI-98 20 D.10 map 82D-82E strain y; cn bw sp, *** SEQUENCING IN PROGRESS***, 73 unordered pieces.	Drosophila melanogaster	35,484	2-Aug-99
rx02610	1050	GB_BA2:AF073776	5138	AF073776	Pseudomonas aeruginosa MexZ (mexZ), complete cds; and mexGH operon, complete sequence.	Pseudomonas aeruginosa	38,323	19-OCT-1998
		GB_BA1:AB015853	5461	AB015853	Pseudomonas aeruginosa gene for MexX and MexY, complete cds.	Pseudomonas aeruginosa	38,447	13-Nov-98
		GB_Vi:HIM237805	2400	AJ237805	Human immunodeficiency virus type 1 partial envelope gene (gp160), isolate MP255.	Human immunodeficiency virus type 1	37,928	15-Jun-99
rx02617	630	GB_PR2:HS181781	4041	L81781	Homo sapiens (subclone Z_e4 from P1 H25) DNA sequence, complete sequence.	Homo sapiens	38,835	9-Apr-97
		GB_HTG3:AC011430	106902	AC011430	Homo sapiens chromosome 5 clone P1_660D11, *** SEQUENCING IN PROGRESS***, 28 unordered pieces.	Homo sapiens	38,662	06-OCT-1999
		GB_HTG3:AC011430	106902	AC011430	Homo sapiens chromosome 5 clone P1_660D11, *** SEQUENCING IN PROGRESS***, 28 unordered pieces.	Homo sapiens	38,662	06-OCT-1999
rx02619	1023	GB_BA1:CG133719	1839	AJ133719	Corynebacterium glutamicum yjc gene, amrR gene and citE gene, partial.	Corynebacterium glutamicum	100,000	12-Aug-99
		GB_BA1:CG133719	1839	AJ133719	Corynebacterium glutamicum yjc gene, amrR gene and citE gene, partial.	Corynebacterium glutamicum	100,000	12-Aug-99
		GB_EST18:AA734344	378	AA734344	w24f02.r1 StrataGene mouse heart (#937316) Mus musculus cDNA clone IMAGE:1223355 5', mRNA sequence.	Mus musculus	37,968	7-Jan-98
rx02620	789	GB_BA1:CG133719	1839	AJ133719	Corynebacterium glutamicum yjc gene, amrR gene and citE gene, partial.	Corynebacterium glutamicum	100,000	12-Aug-99
		GB_BA1:CG133719	1839	AJ133719	Corynebacterium glutamicum yjc gene, amrR gene and citE gene, partial.	Corynebacterium glutamicum	100,000	12-Aug-99
		GB_EST16:C27457	252	C27457	C27457 Rice callus cDNA Oryza sativa cDNA clone C51917.1A, mRNA sequence.	Oryza sativa	39,044	6-Aug-97
rx02624	1419	GB_HTG2:AC005995	170023	AC005995	Homo sapiens clone D10042M02, *** SEQUENCING IN PROGRESS***, 13 unordered pieces.	Homo sapiens	38,533	23-Apr-99
		GB_HTG2:AC005995	170023	AC005995	Homo sapiens clone D10042M02, *** SEQUENCING IN PROGRESS***, 13 unordered pieces.	Homo sapiens	38,533	23-Apr-99
		GB_PR4:AC004895	152927	AC004895	Homo sapiens clone D0810E06, complete sequence.	Homo sapiens	38,533	29-MAY-1999
rx02639	957	GB_BA1:MSGY409	41321	AD000017	Mycobacterium tuberculosis sequence from clone y409.	Mycobacterium tuberculosis	50,370	10-DEC-1996
		GB_BA1:MTCY409	30352	Z97188	Mycobacterium tuberculosis H37Rv complete genome; segment 158/162.	Mycobacterium tuberculosis	38,039	17-Jun-98
		GB_HTG2:AC006765	274498	AC006765	Caenorhabditis elegans clone Y43H11, *** SEQUENCING IN PROGRESS***, 7 unordered pieces.	Caenorhabditis elegans	35,225	23-Feb-99
rx02647	585	GB_PAT:A45577	1925	A45577	Sequence 1 from Patent WO95/19442.	Corynebacterium glutamicum	37,545	07-MAR-1997
		GB_PAT:A45579	1925	A45579	Sequence 3 from Patent WO95/19442.	Corynebacterium glutamicum	37,545	07-MAR-1997
		GB_PAT:A45581	1925	A45581	Sequence 5 from Patent WO95/19442.	Corynebacterium glutamicum	37,545	07-MAR-1997
rx02649	1125	GB_BA1:D90904	150894	D90904	Synechocystis sp. PCC6803 complete genome, 6/27, 630555-781448.	Synechocystis sp.	37,772	7-Feb-99
		GB_PR3:HS941F9	127587	Z95331	Human DNA sequence from BAC 941F9 on chromosome 22q11.2-qter. Contains ESTs, STSs and 3' part of FIBULIN-1 D PRECURSOR like gene, part of a Brain Protein E46 like gene and a CpG island, complete sequence.	Homo sapiens	38,748	23-Nov-99
		GB_BA1:D90904	150894	D90904	Synechocystis sp. PCC6803 complete genome, 6/27, 630555-781448.	Synechocystis sp.	36,364	7-Feb-99
rx02652	702	GB_BA2:AF048749	24454	AF048749	Bacteroides fragilis capsular polysaccharide biosynthesis operon, complete sequence.	Bacteroides fragilis	37,464	25-Jun-99
		GB_HTG4:AC009244	318775	AC009244	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS***, 36 unordered pieces.	Homo sapiens	35,447	26-OCT-1999

TABLE 4: ALIGNMENT RESULTS

rx02655	1107	GB_HTG4:AC009244	318775	AC009244	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS ***; 36 unordered pieces.	Homo sapiens	35,447	26-OCT-1999
		GB_EST10:AA185689	290	AA185689	mi62d11.1 Soares 2NDMT Mus musculus cDNA clone IMAGE 634485 5' similar to TR/G300372 G300372 CELL GROWTH REGULATING NUCLEOLAR PROTEIN, mRNA sequence.	Mus musculus	39,510	19-Feb-97
		GB_GSS13:QA429764	439	QA429764	HS_5079_A2_C12_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=655 Col=24 Row=E, genomic survey sequence.	Homo sapiens	37,701	31-MAR-1999
		GB_GSS8:QA012380	646	QA012380	CIT-HSP-2300L9.TF CIT-HSP Homo sapiens genomic clone 2300L9, genomic survey sequence.	Homo sapiens	38,874	6-Jun-98
rx02662	387	GB_HTG2:AC007802	118569	AC007802	Drosophila melanogaster chromosome 2 clone BACR07111 (D648) RPC1-98 07.1.11 map 58A1-58A2 strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Drosophila melanogaster	37,696	2-Aug-99
		GB_HTG2:AC007802	118569	AC007802	Drosophila melanogaster chromosome 2 clone BACR07111 (D648) RPC1-98 07.1.11 map 58A1-58A2 strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Drosophila melanogaster	37,696	2-Aug-99
		GB_HTG2:AC007802	118569	AC007802	Drosophila melanogaster chromosome 2 clone BACR07111 (D648) RPC1-98 07.1.11 map 58A1-58A2 strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Drosophila melanogaster	35,616	2-Aug-99
rx02665	603	GB_HTG1:AC002419	128340	AC002419	Homo sapiens chromosome X clone BWXD40, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	32,941	12-Aug-97
		GB_HTG1:AC002419	128340	AC002419	Homo sapiens chromosome X clone BWXD40, *** SEQUENCING IN PROGRESS ***; 2 unordered pieces.	Homo sapiens	32,941	12-Aug-97
		GB_GSS15:QA655091	630	QA655091	Sheared DNA-21B5.TF Sheared DNA Trypanosoma brucei genomic clone Sheared DNA-21B5, genomic survey sequence.	Trypanosoma brucei	41,794	22-Jun-99
rx02670	705	GB_HTG2:AC006450	177555	AC006450	Homo sapiens chromosome 9 clone hRPK.85_O_21 map 9, *** SEQUENCING IN PROGRESS ***; 2 ordered pieces.	Homo sapiens	41,374	15-Jul-99
		GB_HTG2:AC006450	177555	AC006450	Homo sapiens chromosome 9 clone hRPK.85_O_21 map 9, *** SEQUENCING IN PROGRESS ***; 2 ordered pieces.	Homo sapiens	41,374	15-Jul-99
		GB_BA1:SPU86147	2268	U86147	Synechococcus PCC7942 UDP-N-acetylmutarinate-alanine ligase (murC) gene, partial cds, UDP-N-acetylglucosaminylglucosamine reductase (murB), and mercuric resistance operon regulatory protein (merR) genes, complete cds.	Synechococcus PCC7942	40,177	15-Feb-97
rx02672	1221	GB_BA1:MTV024	8189	AL022075	Mycobacterium tuberculosis H37Rv complete genome, segment 151/162.	Mycobacterium tuberculosis	53,197	17-Jun-98
		GB_BA1:MSGY23	40806	AD000016	Mycobacterium tuberculosis sequence from clone y23.	Mycobacterium tuberculosis	37,500	10-DEC-1996
		GB_EST38:AW029724	634	AW029724	EST272979 tomato callus, TAMU Lycopersicon esculentum cDNA clone cLEC2817 similar to beta-ketoacyl-ACP synthase, putative, mRNA sequence.	Lycopersicon esculentum	40,000	15-Sep-99
rx02673	756	GB_GSS12:QA356661	518	QA356661	CITBI-E1-2531G11.TF CITBI-E1 Homo sapiens genomic clone 2531G11, genomic survey sequence.	Homo sapiens	33,022	24-Jan-99
		GB_IN1:CELC50A2	36582	AF036689	Caenorhabditis elegans cosmid C50A2.	Caenorhabditis elegans	35,443	05-DEC-1997
		GB_HTG2:AC006702	299864	AC006702	Caenorhabditis elegans clone Y104H12b, *** SEQUENCING IN PROGRESS ***; 3 unordered pieces.	Caenorhabditis elegans	35,443	23-Feb-99
rx02678	1041	GB_HTG1:AP000452	83103	AP000452	Homo sapiens chromosome 11 clone PAC2 map 11q11, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	35,098	13-Sep-99
		GB_HTG1:AP000452	83103	AP000452	Homo sapiens chromosome 11 clone PAC2 map 11q11, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	35,098	13-Sep-99
		GB_PR3:HS498124	58190	AL031057	Human DNA sequence from clone 498124 on chromosome 6p22.1-22.3 Contains STS, GSS and a CpG island, complete sequence.	Homo sapiens	35,328	23-Nov-99
rx02680	1407	GB_BA1:MLCB268	38859	AL022602	Mycobacterium leprae cosmid B268.	Mycobacterium leprae	37,797	27-Aug-99
		GB_BA1:MSGB1554CS	36548	L78814	Mycobacterium leprae cosmid B1554 DNA sequence.	Mycobacterium leprae	36,900	15-Jun-96
rx02683		GB_BA1:MSGB1551CS	36548	L78813	Mycobacterium leprae cosmid B1551 DNA sequence.	Mycobacterium leprae	36,900	15-Jun-96
rx02685	780	GB_HTG3:AC008640	198483	AC008640	Homo sapiens chromosome 5 clone CIT978SKB_17P2, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Homo sapiens	37,299	3-Aug-99

TABLE 4: ALIGNMENT RESULTS

GB_HTG3:AC008640	198483	AC008640	Homo sapiens chromosome 5 clone CIT978SKB_17P2, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Homo sapiens	37,299	3-Aug-99
GB_BA2:AE000945	11317	AE000945	Archaeoglobus fulgidus section 162 of 172 of the complete genome.	Archaeoglobus fulgidus	38,760	15-DEC-1997
GB_BA1:CORPHEA	1088	M13774	C. glutamicum pheA gene encoding prephenate dehydratase, complete cds.	Corynebacterium glutamicum	100,000	26-Apr-93
GB_HTG2:AC007554	167095	AC007554	Homo sapiens clone hRPK.74_A_1, *** SEQUENCING IN PROGRESS ***; 10 unordered pieces.	Homo sapiens	35,226	14-MAY-1999
GB_HTG2:AC007554	167095	AC007554	Homo sapiens clone hRPK.74_A_1, *** SEQUENCING IN PROGRESS ***; 10 unordered pieces.	Homo sapiens	35,226	14-MAY-1999
GB_GSS9:AQ110873	414	AQ110873	CIT-HSP-2378K10. TR CIT-HSP Homo sapiens genomic clone 2378K10, genomic survey sequence.	Homo sapiens	34,768	29-Aug-98
GB_OV:DUKFASA	9138	M21635	Duck (A. platyrhynchos) S-acyl fatty acid synthase thioesterase gene, complete cds.	Anas platyrhynchos	39,177	28-Apr-93
GB_OV:DUKFASA	9138	M21635	Duck (A. platyrhynchos) S-acyl fatty acid synthase thioesterase gene, complete cds.	Anas platyrhynchos	35,022	28-Apr-93
GB_HTG4:AC010096	223607	AC010096	Homo sapiens chromosome unknown clone NH0364A16, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	39,301	29-OCT-1999
GB_HTG4:AC010096	223607	AC010096	Homo sapiens chromosome unknown clone NH0364A16, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	39,301	29-OCT-1999
GB_BA2:AF132788	1411	AF132788	Fischerella muscicola small subunit ribosomal RNA gene, partial sequence.	Fischerella muscicola	36,156	1-Sep-99
GB_GSS5:AQ784372	542	AQ784372	HS_3181_A2_H12_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3181 Col=24 Row=O, genomic survey sequence.	Homo sapiens	41,546	3-Aug-99
GB_HTG4:AC010573	80511	AC010573	Drosophila melanogaster chromosome 3L70C12 clone RPC198-2M13, *** SEQUENCING IN PROGRESS ***; 45 unordered pieces.	Drosophila melanogaster	36,839	16-OCT-1999
GB_EST18:AA696238	840	AA696238	GM05354.5 prime GM Drosophila melanogaster ovary BlueScript Drosophila melanogaster cDNA clone GM05354.5 prime, mRNA sequence.	Drosophila melanogaster	39,523	28-Nov-98
GB_BA1:AB003132	4116	AB003132	Corynebacterium glutamicum gene for MurC, FlsQ, FlsZ, complete cds.	Corynebacterium glutamicum	100,000	4-Aug-97
GB_PR4:AC006024	90563	AC006024	Homo sapiens PAC clone DJ1166G19 from 7p12-p11.2, complete sequence.	Homo sapiens	36,049	20-Feb-99
GB_HTG1:HSJ657D12	315458	AL109943	Homo sapiens chromosome X clone RP4-657D12 map q22.1-24, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	36,475	24-Nov-99
GB_EST20:AA874010	365	AA874010	vw87b05.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone IMAGE:1261905 5', mRNA sequence.	Mus musculus	39,452	19-MAR-1998
GB_EST30:AI658486	599	AI658486	tu17a09.x1 NCI_COAP_P28 Homo sapiens cDNA clone IMAGE:2251288 3', mRNA sequence.	Homo sapiens	37,324	10-MAY-1999
GB_EST20:AA874010	365	AA874010	vw87b05.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone IMAGE:1261905 5', mRNA sequence.	Mus musculus	39,452	19-MAR-1998
GB_BA1:SCF85	21420	AL110470	Streptomyces coelicolor cosmid F85.	Streptomyces coelicolor A3(2)	39,497	1-Sep-99
GB_HTG5:AC011629	69344	AC011629	Homo sapiens chromosome 15 clone 334_M_8 map 15, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	37,462	13-Nov-99
GB_HTG3:AC010106	157581	AC010106	Homo sapiens clone NH0575J05, *** SEQUENCING IN PROGRESS ***; 1 unordered pieces.	Homo sapiens	36,220	11-Sep-99
GB_BA1:MTIC65	34331	Z95584	Mycobacterium tuberculosis H37Rv complete genome; segment 50/162.	Mycobacterium tuberculosis	41,867	17-Jun-98
GB_BA1:MSGY348	40056	AD000020	Mycobacterium tuberculosis sequence from clone y348.	Mycobacterium tuberculosis	43,338	10-DEC-1996
GB_PR2:AP000117	151516	AP000117	Homo sapiens genomic DNA of 21q22.1, GART and AML related, Q78C10-149C3 region, segment 20/20, complete sequence.	Homo sapiens	36,037	25-Sep-99
GB_BA2:AF124600	4115	AF124600	Corynebacterium glutamicum chorismate synthase (aroC), shikimate kinase (aroK), and 3-dehydroquinase synthase (aroB) genes, complete cds; and putative cytoplasmic peptidase (pepQ) gene, partial cds.	Corynebacterium glutamicum	37,294	04-MAY-1999
GB_BA2:AF053071	1083	AF053071	Corynebacterium glutamicum dehydroquinase synthetase (aroB) gene, complete cds.	Corynebacterium glutamicum	39,675	12-Sep-98
GB_BA1:BSTHRZ	19861	Z80360	B. subtilis thrZ downstream chromosomal region.	Bacillus subtilis	34,002	24-Jun-98
GB_PL2:ATF9F13	109936	AL080253	Arabidopsis thaliana DNA chromosome 4, BAC clone F9F13 (ESSA project).	Arabidopsis thaliana	39,252	16-Aug-99



TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

GB_PL2:ATT29H11	87011	AL049659	Arabidopsis thaliana DNA chromosome 3, BAC clone T29H11.	Arabidopsis thaliana	34,249	9-Jun-99
GB_EST4:H37460	489	H37460	15589 Lambda-PRL2 Arabidopsis thaliana cDNA clone 181N1777, mRNA sequence.	Arabidopsis thaliana	45,869	30-DEC-1997
GB_GSS13:AQ454067	534	AQ454067	HS_5176_B1_A09_T7A RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=752 Col=17 Row=B, genomic survey sequence.	Homo sapiens	40,854	21-Apr-99
GB_PR3:AC005495	185254	AC005495	Homo sapiens chromosome 17, clone hRPK 293_K_20, complete sequence.	Homo sapiens	33,200	30-OCT-1998
GB_PR3:AC005495	185254	AC005495	Homo sapiens chromosome 17, clone hRPK 293_K_20, complete sequence.	Homo sapiens	36,273	30-OCT-1998
GB_BA1:AB003132	4116	AB003132	Corynebacterium glutamicum gene for MurC, FtsQ, FtsZ, complete cds.	Corynebacterium glutamicum	98,788	4-Aug-97
GB_BA1:BLFTSZ	5546	Y08964	B.lactofermentum murC, ftsQ or divD & ftsZ genes.	Corynebacterium glutamicum	100,000	08-OCT-1998
GB_HTG4:AC009764	213581	AC009764	Homo sapiens chromosome 11 clone 381_O_22 map 11, *** SEQUENCING IN PROGRESS *** 40 unordered pieces.	Homo sapiens	35,504	21-OCT-1999
GB_BA1:BLFTSZ	5546	Y08964	B.lactofermentum murC, ftsQ or divD & ftsZ genes.	Corynebacterium glutamicum	98,606	08-OCT-1998
GB_BA1:AB003132	4116	AB003132	Corynebacterium glutamicum gene for MurC, FtsQ, FtsZ, complete cds.	Corynebacterium glutamicum	97,561	4-Aug-97
GB_BA1:SCI51	40745	AL109848	Streptomyces coelicolor cosmid I51.	Streptomyces coelicolor A3(2)	39,110	16-Aug-99
GB_BA1:CGU43535	2531	U43535	Corynebacterium glutamicum multidrug resistance protein (cmr) gene, complete cds.	Corynebacterium glutamicum	42,043	9-Apr-97
GB_EST1:D40448	361	D40448	RICS2437A Rice shoot Oryza sativa cDNA, mRNA sequence.	Oryza sativa	41,944	11-Nov-94
GB_BA2:AF038430	8330	AF038430	Thiobacillus neapolitanus carboxysome operon, complete sequence.	Thiobacillus neapolitanus	39,202	21-Aug-98
GB_BA1:MTCY270	37586	Z95388	Mycobacterium tuberculosis H37Rv complete genome; segment 96/162.	Mycobacterium tuberculosis	37,427	10-Feb-99
GB_BA1:MTAG84	1458	X77129	M.tuberculosis Ag84 (CIE) gene.	Mycobacterium tuberculosis	49,339	21-Jul-95
GB_BA1:AP000064	247695	AP000064	Aeropyrum pernix genomic DNA, section 7/7.	Aeropyrum pernix	38,790	22-Jun-99
GB_BA1:TVCP	1923	X56901	T.vulgaris cpT gene for carboxypeptidase T.	Thermoactinomyces vulgaris	37,688	30-OCT-1991
GB_GSS4:AQ682162	355	AQ682162	HS_5503_B2_C11_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=1079 Col=22 Row=F, genomic survey sequence.	Homo sapiens	35,294	28-Jun-99
GB_EST21:AA915356	456	AA915356	v229a08.r1 Soares 2NbtMT Mus musculus cDNA clone IMAGE:1327862 5', mRNA sequence.	Mus musculus	42,609	14-Apr-98
GB_PAT:E13655	2260	E13655	gDNA encoding glucose-6-phosphate dehydrogenase.	Corynebacterium glutamicum	99,620	24-Jun-98
GB_GSS13:AQ447106	711	AQ447106	mgxb0004P18f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0004P18f, genomic survey sequence.	Magnaporthe grisea	38,158	8-Apr-99
GB_EST36:AI900850	670	AI900850	sb95c03.y1 Gm-c1012 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1012-413 5' similar to SW1F4Z_TOBAC Q04688 EUKARYOTIC INITIATION FACTOR 4A-15, mRNA sequence.	Glycine max	37,479	06-DEC-1999
GB_PR1:AB026898	270000	AB026898	Homo sapiens DNA, DLEC1 to ORCTL4 gene region, section 1/2 (DLEC1, ORCTL3, ORCTL4 genes, complete cds).	Homo sapiens	39,364	15-MAY-1999
GB_PR1:AB026898	270000	AB026898	Homo sapiens DNA, DLEC1 to ORCTL4 gene region, section 1/2 (DLEC1, ORCTL3, ORCTL4 genes, complete cds).	Homo sapiens	36,275	15-MAY-1999
GB_PR2:AP000498	100000	AP000498	Homo sapiens genomic DNA, chromosome 3p21.3, clone:603 to 320, anti-oncogene region, section 1/3.	Homo sapiens	39,364	28-Sep-99
GB_HTG4:AC010127	189552	AC010127	Homo sapiens chromosome unknown clone NH0002108, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	36,149	29-OCT-1999
GB_HTG4:AC010127	189552	AC010127	Homo sapiens chromosome unknown clone NH0002108, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	36,149	29-OCT-1999
GB_HTG4:AC010127	189552	AC010127	Homo sapiens chromosome unknown clone NH0002108, WORKING DRAFT SEQUENCE, in unordered pieces.	Homo sapiens	35,924	29-OCT-1999

TABLE 4: ALIGNMENT RESULTS

rx02756	1014	GB_RO:CRUGIPR	2785	D38103	Hamster mRNA for GIP (gastric inhibitory polypeptide) receptor, complete cds.	Cricetulus sp.	38,501	8-Feb-99
		GB_HTG1:CNS01DRV	224372	AL118558	Homo sapiens chromosome 14 clone R-1017G21, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,667	15-OCT-1999
		GB_HTG1:CNS01DRV	224372	AL118558	Homo sapiens chromosome 14 clone R-1017G21, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,667	15-OCT-1999
rx02757	669	GB_RO:RNU35245	2051	U35245	Rat vacuolar protein sorting homolog r-vps33b mRNA, complete cds.	Rattus norvegicus	39,058	15-Jan-97
		GB_RO:RNU35245	2051	U35245	Rat vacuolar protein sorting homolog r-vps33b mRNA, complete cds.	Rattus norvegicus	37,481	15-Jan-97
rx02765	882	GB_EST4:H46883	459	H46883	yo19f06.r1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone IMAGE:178403 5' mRNA sequence.	Homo sapiens	40,132	31-Jul-95
		GB_PL1:NTA6235	3931	AJ006235	Nicotiana tabacum DNA fragment for K-alpha right T-DNA border.	Nicotiana tabacum	34,247	6-Feb-99
		GB_HTG3:AC011519	57287	AC011519	Homo sapiens chromosome 19 clone LLNL-F_192H5, *** SEQUENCING IN PROGRESS ***; 36 unordered pieces.	Homo sapiens	38,497	07-OCT-1999
rx02766	516	GB_GSS13:AQ489419	554	AQ489419	RPC1-11-246A3.TV RPC1-11 Homo sapiens genomic clone RPC1-11-246A3, genomic survey sequence.	Homo sapiens	37,821	24-Apr-99
		GB_PR3:HS90L6	190837	Z97353	Human DNA sequence from clone 90L6 on chromosome 22q11.21-11.23. Contains an RPL15 (60S Ribosomal Protein L15) pseudogene, ESTs, STSs and GSSs, complete sequence.	Homo sapiens	38,160	23-Nov-99
		GB_BA1:MTCY373	35516	Z73419	Mycobacterium tuberculosis H37Rv complete genome, segment 57/162.	Mycobacterium tuberculosis	36,701	17-Jun-98
rx02770	1689	GB_BA2:AF038651	4077	AF038651	Corynebacterium glutamicum dipeptide-binding protein (dcbAE) gene, partial cds; adenine phosphoribosyltransferase (apt) and GTP pyrophosphokinase (rel) genes, complete cds; and unknown gene.	Corynebacterium glutamicum	100,000	14-Sep-98
		GB_OV:AF061275	6756	AF061275	Ictalurus punctatus estrogen receptor type alpha mRNA, complete cds.	Ictalurus punctatus	37,117	31-OCT-1998
		GB_RO:MMSCSELN4	3840	X61756	M.musculus rearranged T-cell receptor beta variable region (Vb17a).	Mus musculus	37,933	03-DEC-1991
rx02774	494	GB_HTG2:AF129408	138685	AF129408	Homo sapiens chromosome 21 clone PAC 31K18 map 21q22.3, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,241	04-MAR-1999
		GB_HTG2:AF129408	138685	AF129408	Homo sapiens chromosome 21 clone PAC 31K18 map 21q22.3, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	38,241	04-MAR-1999
		GB_PR3:AF064859	122571	AF064859	Homo sapiens chromosome 21q22.3 PAC 141B3, complete sequence, containing ribosomal protein homologue pseudogene L23a.	Homo sapiens	37,910	2-Jun-98
rx02775	348	GB_PR3:HSU19F10	31474	Z81145	Human DNA sequence from cosmid U19F10, between markers DXS366 and DXS87 on chromosome X contains ESTs.	Homo sapiens	36,443	23-Nov-99
		GB_EST11:AA211850	293	AA211850	zr90e04.s1 NCL_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:682974 3' similar to TR:G1224065 G1224065 MER37. TRANSPOSABLE ELEMENT, COMPLETE CONSENSUS SEQUENCE. ; mRNA sequence.	Homo sapiens	30,479	13-Aug-97
rx02776	465	GB_EST13:AA372532	400	AA372532	EST84441 Colon adenocarcinoma IV Homo sapiens cDNA 5' end, mRNA sequence.	Homo sapiens	40,104	21-Apr-97
		GB_GSS9:AQ096114	383	AQ096114	HS_3030_A2_D11_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plates-3030 Col=22 Row=G, genomic survey sequence.	Homo sapiens	44,643	27-Aug-98
		GB_RO:AF139518	8827	AF139518	Rattus norvegicus A-kinase anchor protein mRNA, complete cds.	Rattus norvegicus	39,722	16-Jun-99
		GB_EST32:A1728295	564	A1728295	BNLGH10386 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (AF076274) contains similarity to rat p47 protein (GB:AB002086) [Arabidopsis thaliana], mRNA sequence.	Gossypium hirsutum	37,049	11-Jun-99
rx02777	1230	GB_HTG3:AC009407	207973	AC009407	Homo sapiens clone NH022310, *** SEQUENCING IN PROGRESS ***; 16 unordered pieces.	Homo sapiens	39,784	21-Aug-99
		GB_HTG3:AC009407	207973	AC009407	Homo sapiens clone NH022310, *** SEQUENCING IN PROGRESS ***; 16 unordered pieces.	Homo sapiens	39,784	21-Aug-99
		GB_BA2:AE001137	44380	AE001137	Borrelia burgdorferi (section 23 of 70) of the complete genome.	Borrelia burgdorferi	39,105	15-DEC-1997
rx02778	348	GB_EST37:A1940900	668	A1940900	sb79d08.y1 Gm-c1010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID Gm-c1010-1192 5' similar to SW:ITRA_SOYBN P01070 TRYPSIN INHIBITORS A AND C PRECURSOR ; mRNA sequence.	Glycine max	35,036	3-Aug-99
		GB_HTG4:AC008940	131864	AC008940	Homo sapiens chromosome 5 clone CITB-H1_2319M24, *** SEQUENCING IN PROGRESS ***; 3 ordered pieces.	Homo sapiens	33,043	31-OCT-1999
		GB_HTG4:AC008940	131864	AC008940	Homo sapiens chromosome 5 clone CITB-H1_2319M24, *** SEQUENCING IN PROGRESS ***; 3 ordered pieces.	Homo sapiens	33,043	31-OCT-1999

TABLE 4: ALIGNMENT RESULTS

rx02779	402	GB_PR1:AB020876	100000	AB020876	Homo sapiens genomic DNA of 9q32 anti-oncogene of flat epithelium cancer, segment 8/10.	Homo sapiens	35,000	21-MAY-1999
		GB_EST26:AI339126	415	AI339126	q06d03.x1 NCI_CGAP_G04 Homo sapiens cDNA clone IMAGE:1946789 3' similar to gb:L25444	Homo sapiens	35,876	29-DEC-1998
					60S RIBOSOMAL PROTEIN L35A (HUMAN), mRNA sequence.			
rx02780	450	GB_PR1:AB020876	100000	AB020876	Homo sapiens genomic DNA of 9q32 anti-oncogene of flat epithelium cancer, segment 8/10.	Homo sapiens	35,878	21-MAY-1999
		GB_GSS14:AC0524732	591	AQ0524732	HS_5236_A1_D09_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone	Homo sapiens	39,726	11-MAY-1999
					Plate=812 Col=17 Row=G, genomic survey sequence.			
		GB_PR4:AC006051	38903	AC006051	Homo sapiens 12p13.3 PAC RPC15-1103G8 (Roswell Park Cancer Institute Human PAC Library)	Homo sapiens	35,068	26-Nov-98
					complete sequence.			
		GB_PR4:AC006051	38903	AC006051	Homo sapiens 12p13.3 PAC RPC15-1103G8 (Roswell Park Cancer Institute Human PAC Library)	Homo sapiens	37,273	26-Nov-98
					complete sequence.			
rx02781	1314	GB_VI:BHV1CGEN	135301	AJ004801	Bovine herpesvirus 1 complete genome.	Bovine herpesvirus type 1	41,053	11-Jan-99
		GB_BA1:SHGCP1R	107379	X86780	S.hydroscopiscus gene cluster for polyketide immunosuppressant rapamycin.	Streptomyces hydroscopiscus	40,551	16-Aug-96
		GB_BA1:SHGCP1R	107379	X86780	S.hydroscopiscus gene cluster for polyketide immunosuppressant rapamycin.	Streptomyces hydroscopiscus	39,598	16-Aug-96
rx02782	1422	GB_IN1:DSV28T23	5218	X60176	D.silvestris clone U28T2 non-LTR retrotransposon DNA (5218 bp).	Drosophila silvestris	38,895	25-MAR-1992
		GB_IN1:DSV28T24	7779	X60177	D.silvestris clone U28T2 non-LTR retrotransposon DNA (7779 bp).	Drosophila silvestris	38,844	30-Nov-93
		GB_BA2:MPU34795	24888	U34795	Mycoplasma pneumoniae cosmid pcosMPG19 25kb EcoRI fragment	Mycoplasma pneumoniae	37,832	06-MAR-1996
rx02783	672	GB_EST2:R13483	481	R13483	yf7g03.r1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:28494 5', mRNA sequence.	Homo sapiens	36,243	12-Apr-95
		GB_HTG3:AC011333	159160	AC011333	Homo sapiens chromosome 5 clone CIT-HSPC_229L21, *** SEQUENCING IN PROGRESS ***; 23 unordered pieces.	Homo sapiens	38,246	06-OCT-1999
		GB_HTG3:AC011333	159160	AC011333	Homo sapiens chromosome 5 clone CIT-HSPC_229L21, *** SEQUENCING IN PROGRESS ***; 23 unordered pieces.	Homo sapiens	38,246	06-OCT-1999
rx02784	309	GB_PR3:HS474A14	107352	AL023285	Human DNA sequence from clone 474A14 on chromosome 1q24.1-25.2 Contains EST, CA repeat, 5'UTR (tenascin-R), GSS, complete sequence.	Homo sapiens	37,868	23-Nov-99
		GB_GSS12:AQ374565	441	AQ374565	RPC11-159P12.TV RPC1-11 Homo sapiens genomic clone RPC1-11-159P12, genomic survey sequence.	Homo sapiens	36,522	20-MAY-1999
		GB_GSS5:AQ780184	686	AQ780184	HS_3138_B1_F11_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3138 Col=21 Row=L, genomic survey sequence.	Homo sapiens	44,186	2-Aug-99
rx02786	830	GB_PL1:SCYBL033C	1794	Z35794	S.cerevisiae chromosome II reading frame ORF YBL033c.	Saccharomyces cerevisiae	39,264	11-MAR-1998
		GB_PL1:ECRIB1GN	1747	Z21617	S.cerevisiae RIB1 gene encoding GTP cyclohydrolase II.	Saccharomyces cerevisiae	37,118	28-Jan-95
		GB_PAT:A38763	1747	A38763	Sequence 1 from Patent WO9411515.	Saccharomyces cerevisiae	37,118	05-MAR-1997
rx02789	669	GB_EST34:AV146372	295	AV146372	AV146372 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA clone 2810453K19, mRNA sequence.	Mus musculus	65,432	3-Jul-99
		GB_GSS5:AQ785226	524	AQ785226	HS_2025_B2_A07_TTC CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2025 Col=14 Row=B, genomic survey sequence.	Homo sapiens	40,342	3-Aug-99
		GB_BA1:SCOE94	38532	AL049628	Streptomyces coelicolor cosmid E94.	Streptomyces coelicolor	35,346	12-Apr-99
rx02793	825	GB_HTG2:AC006514	247029	AC006514	Homo sapiens clone RPC111-656E20, *** SEQUENCING IN PROGRESS ***; 90 unordered pieces.	Homo sapiens	33,742	1-Apr-99
		GB_HTG2:AC006514	247029	AC006514	Homo sapiens clone RPC111-656E20, *** SEQUENCING IN PROGRESS ***; 90 unordered pieces.	Homo sapiens	33,742	1-Apr-99
		GB_EST22:A033150	384	A033150	ow94b01.s1 Soares fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:1654441 3' similar to TR:O02711 O02711 PRO-POL-DUTPASE POLYPROTEIN, mRNA sequence.	Homo sapiens	38,951	28-Aug-98
rx02796	672	GB_OV:AF172144	1101	AF172144	Xenopus laevis potassium channel beta 2 subunit mRNA, partial cds.	Xenopus laevis	40,000	25-Sep-99
		GB_HTG1:AP000572	84439	AP000572	Homo sapiens chromosome 11 clone p2802 map 11q13, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	35,725	09-OCT-1999

TABLE 4: ALIGNMENT RESULTS

rx02798	1548	GB_HTG1:AP000572	84439	AP000572	Homo sapiens chromosome 11 clone P28D2 map 11q13, *** SEQUENCING IN PROGRESS ***; in Homo sapiens unordered pieces.	35,725	09-OCT-1999
		GB_EST10:AA155097	322	AA155097	mr24h12.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:598439 5', mRNA sequence. Mus musculus	41,195	16-Feb-97
		GB_GSS9:AA150317	489	AQ150317	HS_3178_B1_F12_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3178 Col=23 Row=L, genomic survey sequence.	39,549	08-OCT-1998
rx02799	396	GB_BA2:AE001731	10186	AE001731	Thermotoga maritima section 43 of 136 of the complete genome.	35,747	2-Jun-99
		GB_IN2:AF139876	1562	AF139876	Giardia intestinalis histone H4 gene, complete cds.	34,848	02-OCT-1999
		GB_HTG5:AC011614	176606	AC011614	Drosophila melanogaster chromosome X clone BACR48L05 (D1142) RPCI-98 48 L.5 map 16F-17A strain Y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 114 unordered pieces.	38,817	19-Nov-99
		GB_GSS10:AQ254726	565	AQ254726	EP(3)3517 Drosophila melanogaster EP line Drosophila melanogaster genomic Sequence recovered from Both 5' and 3' ends of P element, genomic survey sequence.	38,798	28-Jun-99
rx02812							
rx02815	552	GB_BA2:PAU32853	1961	U32853	Pseudomonas aeruginosa mucC and mucD genes, complete cds.	39,925	06-MAR-1996
		GB_BA2:PAU49151	4587	U49151	Pseudomonas aeruginosa alternate sigma factor (algU), mucA, mucB, mucC and mucD genes, complete cds.	39,925	13-MAR-1996
		GB_BA2:PAU49151	4587	U49151	Pseudomonas aeruginosa alternate sigma factor (algU), mucA, mucB, mucC and mucD genes, complete cds.	35,922	13-MAR-1996
rx02817	499	GB_BA1:CORAIA	4705	L09232	Corynebacterium glutamicum acetohydroxy acid synthase (ilvB) and (ilvN) genes, and acetohydroxy acid isomerase (ilvC) gene, complete cds.	38,554	23-Feb-95
		GB_EST22:AI004214	475	AI004214	q94h01.x1 Soares_tota_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:1624465 3', mRNA sequence.	51,092	27-Aug-98
		GB_EST27:AI457904	455	AI457904	ij48b08.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2144727 3', mRNA sequence.	37,381	13-Apr-99
rx02818	606	GB_PL1:MZEAHDH1CM	6167	M32984	Zea mays alcohol dehydrogenase (ADH-1 C-m allele) gene, complete cds.	33,111	27-Apr-93
		GB_PL1:ZMADH1ALL	6158	X17556	Z. mays DNA for Adh1-Cm allele.	33,111	23-Jan-92
		GB_PL1:MZEAHDH1CM	6167	M32984	Z. mays alcohol dehydrogenase (ADH-1 C-m allele) gene, complete cds.	38,065	27-Apr-93
rx02823	370	GB_PL2:AF077130	4392	AF077130	Oryza sativa receptor-like protein kinase gene, complete cds.	33,243	25-Jul-98
		GB_HTG1:CNS00M8S	214599	AL079302	Homo sapiens chromosome 14 clone R-1089B7, *** SEQUENCING IN PROGRESS ***; in ordered pieces.	38,420	15-OCT-1999
		GB_HTG1:CNS00M8S	214599	AL079302	Homo sapiens chromosome 14 clone R-1089B7, *** SEQUENCING IN PROGRESS ***; in ordered pieces.	38,420	15-OCT-1999
rx02825	1962	GB_PL1:ZMU17350	1289	U17350	Zea mays thiamine biosynthetic enzyme (thi1-1) mRNA, complete cds.	40,710	12-Apr-96
		GB_GSS13:AA440082	432	AQ440082	HS_5080_B1_B12_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=656 Col=23 Row=D, genomic survey sequence.	41,121	31-MAR-1999
		GB_EST10:AA141278	579	AA141278	CK01542.3prime CK Drosophila melanogaster embryo BlueScript Drosophila melanogaster cDNA clone CK01542.3prime, mRNA sequence.	42,424	29-Nov-98
rx02827	523	GB_BA2:AF054624	6984	AF054624	Lactobacillus sakei transcription-repair coupling factor (trfA) gene, partial cds; L-lactate dehydrogenase (ldhL) gene, complete cds; and unknown genes.	37,241	12-Jan-99
		GB_BA2:AF023181	6616	AF023181	Listeria monocytogenes transcription-repair coupling factor (trfA), low temperature requirement B protein (lfbB), and DivIC homolog (divL) genes, complete cds.	61,228	1-Jan-99
		GB_HTG3:AC009213	114735	AC009213	Drosophila melanogaster chromosome 3 clone BACR09F18 (D812) RPCI-98 09 F.18 map 98D-98D strain Y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 109 unordered pieces.	32,427	23-Aug-99
rx02838	528	GB_BA2:AE000089	12230	AE000089	Rhizobium sp. NGR234 plasmid pNGR234a, section 6 of 46 of the complete plasmid sequence.	57,803	12-DEC-1997
		GB_OM:SSU19994	2004	U19994	Sus scrofa p55 TNF receptor mRNA, complete cds.	34,848	30-Jan-96
		GB_BA2:AE000089	12230	AE000089	Rhizobium sp. NGR234 plasmid pNGR234a, section 6 of 46 of the complete plasmid sequence.	36,505	12-DEC-1997
rx02840	273	GB_BA2:ECU73857	128824	U73857	Escherichia coli chromosome minutes 6-8.	100,000	14-Jul-99
		GB_BA2:AE000145	11448	AE000145	Escherichia coli K-12 MG1655 section 35 of 400 of the complete genome.	100,000	12-Nov-98
		GB_BA1:ECORFABC	3044	X76979	E. coli orf302, orf303 and orf101 sequence.	40,672	28-Aug-96

TABLE 4: ALIGNMENT RESULTS

TABLE 4: ALIGNMENT RESULTS

rx02841	558	GB_GSS4:AQ724171	407	AQ724171	HS_5380_B1_H09_SP6 RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=956 Col=17 Row=P, genomic survey sequence.	Homo sapiens	35,135	14-Jul-99
		GB_PL1:PSFERR	1023	X64417	P.sativum mRNA for ferritin.	Pisum sativum	41,985	30-Jun-93
		GB_PL1:PSFERRI	1023	X73369	P.sativum mRNA for ferritin.	Pisum sativum	41,985	21-Jul-95
rx02842	365	GB_EST32:A1723928	435	A1723928	RHIZ1_35_B10.y1_A001 Rhizome1 Sorghum halepense cDNA, mRNA sequence.	Sorghum halepense	32,616	11-Jun-99
		GB_PR2:HS212A2	212753	Z95114	Human DNA sequence from clone 212A2 on chromosome 22q12 Contains gene for TNF-inducible protein CG12-1, 3' end of a gene similar to apolipoprotein L, ESTs, STSs, CA repeat, GSSs and CpG islands, complete sequence.	Homo sapiens	38,279	23-Nov-99
		GB_OV:AF116856	4634	AF116856	Gallus gallus neurocan core protein precursor, mRNA, complete cds.	Gallus gallus	41,176	20-MAY-1999
rx02845	615	GB_PR3:AC005934	43699	AC005934	Homo sapiens chromosome 19, cosmid R30813, complete sequence.	Homo sapiens	35,225	5-Nov-98
		GB_PR3:AC005340	36705	AC005340	Homo sapiens chromosome 19, cosmid F19544, complete sequence.	Homo sapiens	35,225	30-Jul-98
		GB_PR3:AC005934	43699	AC005934	Homo sapiens chromosome 19, cosmid R30813, complete sequence.	Homo sapiens	39,153	5-Nov-98
rx02846	873	GB_GSS5:AQ749564	838	AQ749564	HS_5575_A1_B11_SP6 RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1151 Col=21 Row=C, genomic survey sequence.	Homo sapiens	36,286	19-Jul-99
		GB_GSS5:AQ749564	838	AQ749564	HS_5575_A1_B11_SP6 RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1151 Col=21 Row=C, genomic survey sequence.	Homo sapiens	38,591	19-Jul-99
rx02847	1135	GB_BA1:MSG813GS	42923	L78823	Mycobacterium leprae cosmid B13 DNA sequence.	Mycobacterium leprae	47,559	15-Jun-96
		GB_HTG4:AC009557	128590	AC009557	Homo sapiens chromosome 15 clone 76_D_16 map 15, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	36,289	25-OCT-1999
		GB_HTG4:AC009557	128590	AC009557	Homo sapiens chromosome 15 clone 76_D_16 map 15, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	36,289	25-OCT-1999
rx02848	499	GB_BA1:AB018531	4961	AB018531	Corynebacterium glutamicum disR1 and disR2 genes, complete cds.	Corynebacterium glutamicum	98,397	19-OCT-1998
		GB_PAT:E17019	4961	E17019	Brevibacterium lactofermentum disR and disR2 genes.	Corynebacterium glutamicum	98,397	28-Jul-99
		GB_OM:CFU08596	1780	U08596	Canis familiaris delayed rectifier K+ channel mRNA, partial cds.	Canis familiaris	43,265	21-DEC-1994
rx02849	305	GB_VI:AF121950	9395	AF121950	Hepatitis G virus strain Iowan, complete genome.	Hepatitis G virus	33,110	24-MAY-1999
		GB_PAT:AR049304	9327	AR049304	Sequence 234 from patent US 5824507.	Unknown.	31,803	29-Sep-99
		GB_PAT:AR026878	9327	AR026878	Sequence 234 from patent US 5856134.	Unknown.	31,803	29-Sep-99